

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 05:53:08 ; Search time 15911 Seconds
(without alignments)
11377.987 Million cell updates/sec

Title: US-10-726-160-1

Perfect score: 2831

Sequence: 1 ggcacgagcggagtggag.....atgaaaaaaaaaaaaaa 2831

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_env.*

2: gb_pat.*

3: gb_ph.*

4: gb_pl.*

5: gb_pr.*

6: gb_ro.*

7: gb_sts.*

8: gb_sy.*

9: gb_un.*

10: gb_vi.*

11: gb_ov.*

12: gb_htg.*

13: gb_in.*

14: gb_om.*

15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2831	100.0	2831	5	AF345896	AF345896 Homo sapi
2	2816.2	99.5	2838	2	AX213281	AX213281 Sequence
3	2804	99.0	4422	2	CQ413115	CQ413115 Sequence
4	2804	99.0	4422	2	CQ493832	CQ493832 Sequence
5	2804	99.0	4422	2	AX188263	AX188263 Sequence
6	2802.4	99.0	3008	5	AK000742	AK000742 Homo sapi
7	2802.2	99.0	2969	5	BC033540	BC033540 Homo sapi
8	2802	99.0	4450	2	AX780410	AX780410 Sequence
9	2789.8	98.5	4221	2	AX926533	AX926533 Sequence
10	2789.8	98.5	4221	2	AX951786	AX951786 Sequence
11	2789.8	98.5	4221	2	AX960064	AX960064 Sequence
12	2789.8	98.5	4221	2	CS033737	CS033737 Sequence
13	2789.8	98.5	4221	2	CS042689	CS042689 Sequence
14	2789.8	98.5	4221	5	AF195765	AF195765 Homo sapi
15	2775.6	98.0	2782	5	BC033297	BC033297 Homo sapi
16	2290.6	80.9	3685	2	BD159852	BD159852 Primer fo
17	2290.6	80.9	3685	2	AX882660	AX882660 Sequence
18	2290.6	80.9	3685	5	AK027651	AK027651 Homo sapi

19	1833	64.7	1833	2	BD156087	BD156087 Primer fo
20	1833	64.7	1833	2	AX876358	AX876358 Sequence
21	1833	64.7	1833	5	AK001261	AK001261 Homo sapi
22	1812.6	64.0	2442	6	AB095735	Mus muscu
23	1783.8	63.0	4160	6	BC060208	Mus muscu
24	1729.2	61.1	3175	6	AB095736	Mus muscu
25	1520	53.7	1871	2	AX274866	Sequence
26	1490.4	52.6	1803	2	CQ728118	Sequence
27	1273.8	45.0	1291	2	BD156017	BD156017 Primer fo
28	1273.8	45.0	1291	2	AX876224	Sequence
29	1273.8	45.0	1291	5	AK001206	Homo sapi
30	856.4	30.3	905	2	BD147262	BD147262 Primer fo
31	856.4	30.3	905	2	AX867200	Sequence
32	843	29.8	3017	11	AJ720316	AJ720316 Gallus ga
33	834.2	29.5	160200	5	AL606468	Human DNA
34	804.4	28.4	853	2	BD147403	BD147403 Primer fo
35	804.4	28.4	853	2	AX867341	Sequence
36	680.6	24.0	215416	12	AC153228	AC153228 Bos tauru
37	680.6	24.0	232073	12	AC171797	AC171797 Bos tauru
38	652.2	23.0	2904	11	BC097560	BC097560 Xenopus l
39	649.8	23.0	205300	6	AC132830	AC132830 Mus muscu
40	649.4	22.9	3639	11	CR848464	CR848464 Xenopus t
41	649.4	22.9	2676	11	BC064851	BC064851 Xenopus t
42	633.6	22.4	2778	11	BC073015	BC073015 Xenopus l
43	625.4	22.1	242922	12	AC105627	AC105627 Rattus no
44	581.6	20.5	1309	11	CR386886	CR386886 Gallus ga
45	558	19.7	708	2	BD129416	BD129416 Human gen

ALIGNMENTS

RESULT 1	AF345896	Homo sapiens RA-regulated nuclear matrix-associated protein mRNA, complete cds.	2831 bp	mRNA	linear	PRI 21-MAY-2001
AF345896	LOCUS	AF345896	2831 bp	mRNA	linear	PRI 21-MAY-2001
DEFINITION	AF345896.1	GI:14161319				
ACCESSION	AF345896					
VERSION	AF345896.1	GI:14161319				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
PUBMED						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
CDS						

1921 AAGGACTCTCTAGGTCCTACCAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCA 1980
1921 AAGGACTCTCTAGGTCCTACCAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCA 1980
1981 GAGCCTCCGTCCTCTATCAGTCGGTATGCTTCAAGAAAGCTGTGGAACGCTACTCTTCTCT 2040
1981 GAGCCTCCGTCCTCTATCAGTCGGTATGCTTCAAGAAAGCTGTGGAACGCTACTCTTCTCT 2040
2041 TTGAGACCTTGTGGAGAAAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAT 2100
2041 TTGAGACCTTGTGGAGAAAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAT 2100
2101 AAAAATCTGTTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAAT 2160
2101 AAAAATCTGTTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAAT 2160
2161 CGGTCAATCCAGACACCAATTCAGGAGACAGAGCGGAAAGACATGTCGAAGCCCGGTC 2220
2161 CGGTCAATCCAGACACCAATTCAGGAGACAGAGCGGAAAGACATGTCGAAGCCCGGTC 2220
2221 ACCATCAGCCCGAGCTCCATGAGGAAATCTGCACATATCTCCATAGAAAGTCCAGAGAG 2280
2221 ACCATCAGCCCGAGCTCCATGAGGAAATCTGCACATATCTCCATAGAAAGTCCAGAGAG 2280
2281 GACTTCTGTGTCCTGAACACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTCA 2340
2281 GACTTCTGTGTCCTGAACACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTCA 2340
2341 GCTTTGGTCCACTAAACAGCTGAGCTTTGGTCCACTAAACAGATGAAATATCAAG 2400
2341 GCTTTGGTCCACTAAACAGCTGAGCTTTGGTCCACTAAACAGATGAAATATCAAG 2400
2401 AGTGACTCTATTAATCTGCTGTTTGAAGAGCTGCCTTTTCAATTTTGAACAAATCTTT 2460
2401 AGTGACTCTATTAATCTGCTGTTTGAAGAGCTGCCTTTTCAATTTTGAACAAATCTTT 2460
2461 TCAACGCTGAAATGACTAAATCTGCTTCTACTACCAATATGATATGAGCTTCCCGAG 2520
2461 TCAACGCTGAAATGACTAAATCTGCTTCTACTACCAATATGATATGATGAGCTTCCCGAG 2520
2521 GATGAATGCTGTTTAAATTTTCAAAAGTAAATTTGTCACTCTAGCAATTTGAATGAAT 2580
2521 GATGAATGCTGTTTAAATTTTCAAAAGTAAATTTGTCACTCTAGCAATTTGAATGAAT 2580
2581 AGTCTTCACTTTTAAATTTTCAATCTCTATATATGATGATGATGATGATGATGATGATGAT 2640
2581 AGTCTTCACTTTTAAATTTTCAATCTCTATATATGATGATGATGATGATGATGATGATGAT 2640
2641 CAAAAACAAAGTTTCTTGTATCTGTAACCTTTCTATGCTGAGTGAAGATATCTGCCAG 2700
2641 CAAAAACAAAGTTTCTTGTATCTGTAACCTTTCTATGCTGAGTGAAGATATCTGCCAG 2700
2701 CCACAGATGAGGCTGTGAAGGCTGATGAGAAATCTCTCTGCTGAAGACCCCTGGTTCT 2760
2701 CCACAGATGAGGCTGTGAAGGCTGATGAGAAATCTCTCTGCTGAAGACCCCTGGTTCT 2760
2761 GTTCTGCTCCACATGATATATTTTATTTGAATACATAATCTTTTCACTATGAAAAA 2820
2761 GTTCTGCTCCACATGATATATTTTATTTGAATACATAATCTTTTCACTATGAAAAA 2820
2821 AAAAAAATAA 2831
2821 AAAAAAATAA 2831

RESULT 2
LOCUS AX213281 2838 bp DNA linear PAT 06-SEP-2001
DEFINITION Sequence 1 from Patent WO0159115.
ACCESSION AX213281
VERSION AX213281.1 GI:15524191
KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Hubert.R.S., Afar,D.E., Challita-Bid,P.M., Paris,M., Levin,E.,
Mitchell,S.C. and Jakobovits,A.
TITLE 83p594: a tissue specific protein highly expressed in prostate
cancer
JOURNAL Patent: WO 0159115-A 1 16-AUG-2001;
Urogenesys, Inc. (US)
FEATURES
source Location/Qualifiers
1..2838
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
130..2322
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC69322.1"
/db_xref="GI:15524192"
/translation="MLFNSALRQPOLGVLBNHSSQYPLQSLLTGCGSGNDHSTYSY
ETGVEVPFGCTFSAPMEHVLAVANEEGFVRLYNTSOSFRKCKFEWMAHNAVF
DLAWPGELKLVTAAGDQFADKFWDKAGELIGTCGHQCSLKSVAFSKPAVFTGG
RDGIMVMDTRCNKKGDFRVNQISGAHNTSDKQTPSKPKKQNSKGLAPSDVFOQS
VTVLFDENTLVSAGAVDGIKWMDLRKNTAYRQEPITASKSFLYPGSLRKLGYSS
LILDSGTSLFANCTDDNIYFNMTHGLATSPVAIFNGHONSTFYVKSLSLSDQFLVS
GSSDEAAYIKVSTPWQPTVLLGHSQEVTSVWCPSDFTKIATCSDNTLKLMLNR
GLEKPGDKLSTVGWASKKESRPGLVTVSSQSTPAKAPRVKCNFNSPSSAAAC
APSCAGDILPLSNPTTFSIKTSPAKARPIINRRGSVSVSPKPKSFKMSINRWTRT
PSSPPIPTPASETKIMSPKALIPVQSKSOAEACSESRNRVRKRLSSCLINSEVQKQ
CVKSCNVTELDGOVENLHLCLCLAGNQEDLSGLGPTKSKLIEGAGTSLSEPPSP
ISPVAESCGTLPLPLRPGRGSEVMGKENSSENKMWLLAMAAARKKAEKNPSPSPSS
QTPNSRRSGKTLPSPTVITPSSMRKICTYFHRKSQEDFCGPEHSTEL"

Query Match 99.5%; Score 2816.2; DB 2; Length 2838;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2821; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 3 CACGAGCGGAGTTGGAGCGGATAAACGATTTGTGTGTGAGAGGCGCAACGTGCGATTTC 62
DB 9 CAGTGGCGGAGTTGGAGCGGATAAACGATTTGTGTGTGAGAGGCGCAACGTGCGATTTC 68
QY 63 TGCTCAATCTTGAGGACATTTCTACGACTTTCTCTCAGTCAGGCTTTTCTCCGACCCCT 122
DB 69 TGCTCAATCTTGAGGACATTTCTACGACTTTCTCTCAGTCAGGCTTTTCTCCGACCCCT 128
QY 123 GATGCTCTTCAATTCGGTGTCCGCGCAGCCCGAGCTTGGCGTCTCAGAGAAATGGATGGTC 182
DB 129 GATGCTCTTCAATTCGGCGCTCCGCGCAGCCCGAGCTTGGCGTCTCAGAGAAATGGATGGTC 188
QY 183 TTCACAAATACCTTTTCAATCCCTTCTCAGTGGTTATCAGTCAGTGGTAAATGATGAACA 242
DB 189 TTCACAAATACCTTTTCAATCCCTTCTCAGTGGTTATCAGTCAGTGGTAAATGATGAACA 248
QY 243 CACTTCTTATGGAGAAACAGGAGTCCCGATTTCTCTTTTGGATGATCTTCTCTCTCGC 302
DB 249 CACTTCTTATGGAGAAACAGGAGTCCCGATTTCTCTTTTGGATGATCTTCTCTCTCGC 308
QY 303 TCCCAATATGGAACATGATCTAGCAGTTGCCAATCAAGAGGCTTTTGTTCGATTCTATAA 362
DB 309 TCCCAATATGGAACATGATCTAGCAGTTGCCAATGAGAGGCTTTTGTTCGATTCTATAA 368
QY 363 CACAGAAATCAAAAGTTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGTCTCACTGGAATGC 422
DB 369 CACAGAAATCAAAAGTTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGTCTCACTGGAATGC 428
QY 423 CGTCTTTCAGCTGGCTGGTCTCGTGGTGAACCTTAACTTGTATACAGCAGCAGGTGATCA 482
DB 429 CGTCTTTCAGCTGGCTGGTCTCGTGGTGAACCTTAACTTGTATACAGCAGCAGGTGATCA 488
QY 483 AACAGCCAAATTTTGGGACGTAAGAGCTGGTGAGCTGATTTGGAAACATGCAAGGTCATCA 542

ORIGIN

|||||
489 AACAGCCAAATTTTGGGACGTAAGCTGGTGAAGCTGATTTGGAAATGCAAGAGGTCATCA 548
QY
543 ATGCAAGCCTCAAGTCAGTGTGGCTTTTCTAAGTTTTCAGAAAGCTGTATTTCTGTACGGGTGG 602
Db
549 ATGCAAGCCTCAAGTCAGTGTGGCTTTTCTAAGTTTTCAGAAAGCTGTATTTCTGTACGGGTGG 608
QY
603 AAGAGATGGCAACATATATGGTCTGGGATACCAAGGTGGCAACAAAAAGATGGGTTTATAG 662
Db
609 AAGAGATGGCAACATATATGGTCTGGGATACCAAGGTGGCAACAAAAAGATGGGTTTATAG 668
QY
663 GCAAGTGAATCAAAATCAGTGGAGCTCACAAATACCTCAGCAGCAAGCAACCCCTTCAAAACC 722
Db
669 GCAAGTGAATCAAAATCAGTGGAGCTCACAAATACCTCAGCAGCAAGCAACCCCTTCAAAACC 728
QY
723 CAAGAAGAAACAGAAATTCAAAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAGGTGTTAC 782
Db
729 CAAGAAGAAACAGAAATTCAAAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAGGTGTTAC 788
QY
783 TGTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAT 842
Db
789 TGTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAT 848
QY
843 CAAAGTATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAAGAACCCATAGCATCAA 902
Db
849 CAAAGTATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAAGAACCCATAGCATCAA 908
QY
903 GTCTTTCTCTGTACCCAGGTAGCAGCACTCGAAAACTTTGGATATCAAGTCTGATTTTGA 962
Db
909 GTCTTTCTCTGTACCCAGGTAGCAGCACTCGAAAACTTTGGATATCAAGTCTGATTTTGA 968
QY
963 TTCCAATGGCTCTACTTTATTTGCTAAATTTGCAAGATGCAAGATACATCTCATGTTTAAAT 1022
Db
969 TTCCAATGGCTCTACTTTATTTGCTAAATTTGCAAGATGCAAGATACATCTCATGTTTAAAT 1028
QY
1023 GACTGGTGTGAAGCTTCTCCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTA 1082
Db
1029 GACTGGTGTGAAGCTTCTCCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTA 1088
QY
1083 TGTAAAAATCCAGCCTTAGTCCAGATGACCCAGTTTTTTAGTCAAGTGGCTCAAGTGATGAAGC 1142
Db
1089 TGTAAAAATCCAGCCTTAGTCCAGATGACCCAGTTTTTTAGTCAAGTGGCTCAAGTGATGAAGC 1148
QY
1143 TGCCTACATATGGAAGGTCTCACACCTTGGCAACTCTTACTGTGCTCTGGGTCAATTC 1202
Db
1149 TGCCTACATATGGAAGGTCTCACACCTTGGCAACTCTTACTGTGCTCTGGGTCAATTC 1208
QY
1203 TCAAGAGGTCAAGTCTGTGCTGGTGTCCATCTGACTTCAAAAAGATTTGCTACCTGTTTC 1262
Db
1209 TCAAGAGGTCAAGTCTGTGCTGGTGTCCATCTGACTTCAAAAAGATTTGCTACCTGTTTC 1268
QY
1263 TGATGACATACATAAATAATCTGGCGCTTGAATAGAGCTTTAGAGGAAACCAAGAGG 1322
Db
1269 TGATGACATACATAAATAATCTGGCGCTTGAATAGAGCTTTAGAGGAAACCAAGAGG 1328
QY
1323 TGATAAATCTTCCAGGTGGGTGGGCTCTCAGAAAGAAAAAGATCAAGACCTGGGCT 1382
Db
1329 TGATAAATCTTCCAGGTGGGTGGGCTCTCAGAAAGAAAAAGATCAAGACCTGGGCT 1388
QY
1383 AGTAAACAGTAAACAGGTAGCCAGGTACTCTCTGCAAGGCCCCAGGGTAAAGTGCAATCC 1442
Db
1389 AGTAAACAGTAAACAGGTAGCCAGGTACTCTCTGCAAGGCCCCAGGGTAAAGTGCAATCC 1448
QY
1443 ATCCAAATCTTCCCGCTACCGCAGCTTGTGCCCCCAAGCTGTGCTGGAGACCTCCCTCT 1502
Db
1449 ATCCAAATCTTCCCGCTACCGCAGCTTGTGCCCCCAAGCTGTGCTGGAGACCTCCCTCT 1508
QY
1503 TCCCTTCAAAATACCTCTAATTAATAACCTCTCTGCAAGGCCCGGTCTCCCAT 1562
Db
1509 TCCCTTCAAAATACCTCTAATTAATAACCTCTCTGCAAGGCCCGGTCTCCCAT 1568
QY
1563 CAAACAGAGAGCTGTGCTCTCTCGTCTCTCCCAAGGCCACCTTCATCTTTCAAGATGTC 1622
|||||

Db
1569 CAACAGAAAGGCTCTGTCTCTCCGTCTCTCCCAAGCCACCTTTCACTTTTCAAGATGTC 1628
QY
1623 GATTAGAAAACTGGGTGACCCGAAACACCTTCTCTCATCAACACCACTCACTCCACCTGCTTC 1682
Db
1629 GATTAGAAAACTGGGTGACCCGAAACACCTTCTCTCATCAACACCACTCACTCCACCTGCTTC 1688
QY
1683 GGAGACCAAGATCATGTCTCCGAAAGCCCTTATTCTCTGTAGCCAGAGTCAATCCA 1742
Db
1689 GGAGACCAAGATCATGTCTCCGAAAGCCCTTATTCTCTGTAGCCAGAGTCAATCCA 1748
QY
1743 AGCAGAGCTTGTCTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT 1802
Db
1749 AGCAGAGCTTGTCTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT 1808
QY
1803 GGAAGTGTGAAAAAAGAGTGTGAAGTGTGAATCTGTGTGACTGAGCTTGTATGGCCA 1862
Db
1809 GGAAGTGTGAAAAAAGAGTGTGAAGTGTGAATCTGTGTGACTGAGCTTGTATGGCCA 1868
QY
1863 AGTTGAAAAATCTTCAATTTGGATCTGTCTGCTCTTGTGTAACCAAGGAAGACCTTAGTAA 1922
Db
1869 AGTTGAAAAATCTTCAATTTGGATCTGTCTGCTCTTGTGTAACCAAGGAAGACCTTAGTAA 1928
QY
1923 GGAATCTCTAGTCTTACCAAAATCAAGCAAAATTTGAAGGAGCTGTGTACCAAGTATCTCAGA 1982
Db
1929 GGAATCTCTAGTCTTACCAAAATCAAGCAAAATTTGAAGGAGCTGTGTACCAAGTATCTCAGA 1988
QY
1983 GCCTTCGGTCTCTATCAGTCCGTATGCTTCAAGAAAGCTGTGGAAAGCTTCTTCTCTTT 2042
Db
1989 GCCTTCGGTCTCTATCAGTCCGTATGCTTCAAGAAAGCTGTGGAAAGCTTCTTCTCTTT 2048
QY
2043 GAGACCTTGTGGAGAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTTCCCCAGAGAAATAA 2102
Db
2049 GAGACCTTGTGGAGAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTTCCCCAGAGAAATAA 2108
QY
2103 AAACTGGTGTGTGGCCATGGCAGCAACCAAGAGGCTGAGAAATCCATCTCCACCAAGTCC 2162
Db
2109 AAACTGGTGTGTGGCCATGGCAGCAACCAAGAGGCTGAGAAATCCATCTCCACCAAGTCC 2168
QY
2163 GTCAATCCAGACACCCAAATTCAGSAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAAC 2222
Db
2169 GTCAATCCAGACACCCAAATTCAGSAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAAC 2228
QY
2223 CATCAAGCCAGCTCCATGAGGAAAACTGCACATACCTTCCATAGAAAAGTCCCAAGGAGA 2282
Db
2229 CATCAAGCCAGCTCCATGAGGAAAACTGCACATACCTTCCATAGAAAAGTCCCAAGGAGA 2288
QY
2283 CTTCTGTGGTCTCTGAACAACATCAACAGAAATTAGATTTCTAATCTGAGTGAGTTACTGAGC 2342
Db
2289 CTTCTGTGGTCTCTGAACAACATCAACAGAAATTAGATTTCTAATCTGAGTGAGTTACTGAGC 2348
QY
2343 TTTGGTCCACTAAAACAAAGCTGAGCTTTGGTCCACTAAAACAAAGTGAAGAAATACAAAG 2402
Db
2349 TTTGGTCCACTAAAACAAAGCTGAGCTTTGGTCCACTAAAACAAAGTGAAGAAATACAAAG 2408
QY
2403 TGACTCTATAACTCTGGTCTTTTAAAGAAAGCTGCCCTTTTCAATTTTAAAGAAAAATCTTTTC 2462
Db
2409 TGACTCTATAACTCTGGTCTTTTAAAGAAAGCTGCCCTTTTCAATTTTAAAGAAAAATCTTTTC 2468
QY
2463 AACGCTGAAAATGTACCTAAATCTGGTCTTCACTACCAATAATGTATATGCAAGTTCGCCAGGA 2522
Db
2469 AACGCTGAAAATGTACCTAAATCTGGTCTTCACTACCAATAATGTATATGCAAGTTCGCCAGGA 2528
QY
2523 TGAATGCTGTGTTTAAATTTTCAATTAAGTAAATTTGTCTACTAGCATTTTGAATGAATAG 2582
Db
2529 TGAATGCTGTGTTTAAATTTTCAATTAAGTAAATTTGTCTACTAGCATTTTGAATGAATAG 2588
QY
2583 TCTTCACTTTTAAATTTTCAATTTTCAATAATGAATCACTCCAGTTCATGAGGAGA 2642
Db
2589 TCTTCACTTTTAAATTTTCAATTTTCAATAATGAATCACTCCAGTTCATGAGGAGA 2648
QY
2643 AAAACCAAGTTCCTTGTATCTCTCAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702
Db
2649 AAAACCAAGTTCCTTGTATCTCTCAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2708
|||||


```
QY 2703 ACAGCATAGGCGCTGTGAAGGCTGACTGAGAAATCCTCTGCTCAGACCCCTGGTTCTGT 2762
|||||
Db 2709 ACAGCATAGGCGCTGTGAAGGCTGACTGAGAAATCCTCTGCTGAGACCCCTGGTTCTGT 2768
|||||
QY 2763 TCTGCCTCAACATGTATATTTTATTGAAATACATAATCTTTTCACTATGAAAAAAA 2822
|||||
Db 2769 TCTGCCTCAACATGTATATTTTATTGAAATACATAATCTTTTCACTAAAAAAA 2828
|||||
QY 2823 AAAAAAAA 2831
|||||
Db 2829 AAAAAAAA 2837
|||||

RESULT 3
CQ413115 4422 bp DNA linear PAT 23-JAN-2004
LOCUS Sequence 20186 from Patent WO0170979.
DEFINITION CQ413115
ACCESSION CQ413115
VERSION CQ413115.1 GI:41320896
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 Lee, J. and Lillie, J.
AUTHORS Genes, compositions, kits, and method for identification,
TITLE assessment, prevention, and therapy of ovarian cancer
JOURNAL Patent: WO 0170979-A 20186 27-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..4422
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

- ORIGIN
Query Match 99.0%; Score 2804; DB 2; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACGAGCGGAGTTGGAGCGATAAGATTTCGTGTGAGAGGCGCAAGTGGCATTTTC 62
|||||
Db 8 CAGTGGCGGAGTTGGAGCGATAAGATTTCGTGTGAGAGGCGCAAGTGGCATTTTC 67
|||||
QY 63 TGCTGAATTTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCT 122
|||||
Db 68 TGCTGAATTTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCT 127
|||||
QY 123 GATGCTCTTCAATTCGGTCTCGGCGAGCCCGAGCTGGCGTCTGAGAAATGGATGGTC 182
|||||
Db 128 GATGCTCTTCAATTCGGTCTCGGCGAGCCCGAGCTGGCGTCTGAGAAATGGATGGTC 187
|||||
QY 183 TTCACAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGTGGTAAATGATGACACA 242
|||||
Db 188 TTCACAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGTGGTAAATGATGACACA 247
|||||
QY 243 CACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTTGGATGTACCTTCTCTCTGC 302
|||||
Db 248 CACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTTGGATGTACCTTCTCTCTGC 307
|||||
QY 303 TCCCAATATGGAACATGTACTAGCAGTTCGCCAATGAAGAAGCTTTGTTTCGATTGTATAA 362
|||||
Db 308 TCCCAATATGGAACATGTACTAGCAGTTCGCCAATGAAGAAGCTTTGTTTCGATTGTATAA 367
|||||
QY 363 CACAGAATCACAAAGTTTCAGAAAGAGTCTTCAAGAATGGATGGCTCACTGGAATGC 422
|||||
Db 368 CACAGAATCACAAAGTTTCAGAAAGAGTCTTCAAGAATGGATGGCTCACTGGAATGC 427
|||||
QY 423 CQTCTTTGACCTGGCTGGGTTCTCTGTTGAACTTAACTTTTACAGCAGCAGGTGATCA 482
|||||
```

```
Db 428 CGTCTTTGACCTGGCTGGGTTCTGTGTAACCTTAAACTTGTGTTACAGCAGCAGGTGATCA 487
QY 483 AACACCCAAATTTTGGGACGTAAAGCTGGTGAAGCTGATTTGGAAACATGCAAAAGGTCAATCA 542
Db 488 AACACCCAAATTTTGGGACGTAAAGCTGGTGAAGCTGATTTGGAAACATGCAAAAGGTCAATCA 547
QY 543 ATGCAAGCTCAAGTCAGTTGCTTTTCTAAAGTTTTCAGAAAAGCTGTATTCTGTACGGGTGG 602
Db 548 ATGCAAGCTCAAGTCAGTTGCTTTTCTAAAGTTTTCAGAAAAGCTGTATTCTGTACGGGTGG 607
QY 603 AAGAGATGGCAACATTTATGGTCTGGGATACCAAGTGTCAACAAAAAGATGGGTTTATAG 662
Db 608 AAGAGATGGCAACATTTATGGTCTGGGATACCAAGTGTCAACAAAAAGATGGGTTTATAG 667
QY 663 GCAGTGAATCAATCAGTGGAGCTCAATACTCTCAGACAGCAAAACCCCTTCAAAACC 722
Db 668 GCAGTGAATCAATCAGTGGAGCTCAATACTCTCAGACAGCAAAACCCCTTCAAAACC 727
QY 723 CAAGAAGAAACAGAAATTCAAAAGGACTTGTCTCTCTGTGGATTTTCCAGCAAGTGTATAC 782
Db 728 CAAGAAGAAACAGAAATTCAAAAGGACTTGTCTCTCTGTGGATTTTCCAGCAAGTGTATAC 787
QY 783 TGTGCTCTCTTTCAAGACGAGAAATACCTTAGTCTCTCAGCAGGAGCTGTGGATGGGATAT 842
Db 788 TGTGCTCTCTTTCAAGACGAGAAATACCTTAGTCTCTCAGCAGGAGCTGTGGATGGGATAT 847
QY 843 CAAGATATGGGATTTACGTAAGAAATATACCTTATTCGACAGCAAAACCCCTTCAAAACC 902
Db 848 CAAGATATGGGATTTTACGTAAGAAATATACCTTATTCGACAGCAAAACCCCTTCAAAACC 907
QY 903 GTCTTTCTCTGTACCCAGCTAGCAGCACTCGAAAACCTCGGATATTCAGTCTGATTTTGA 962
Db 908 GTCTTTCTCTGTACCCAGCTAGCAGCACTCGAAAACCTCGGATATTCAGTCTGATTTTGA 967
QY 963 TTCCACTGGCTCTACTTTATTTGTTAAATTCACAGACGATCAATCAATGTTTAAATAT 1022
Db 968 TTCCACTGGCTCTACTTTATTTGCTTAAATTCACAGACGATCAATCAATGTTTAAATAT 1027
QY 1023 GACTGGGTTGAAGACTTCTCAGTGGCTATTTTCAATGGACACCAAGACTTACCTTTTA 1082
Db 1028 GACTGGGTTGAAGACTTCTCAGTGGCTATTTTCAATGGACACCAAGACTTACCTTTTA 1087
QY 1083 TGTAAAATTCAGCCCTTAGTCCAGATGACAGTCTTTTGTAGTGGCTCAAGTGAATGAAGC 1142
Db 1088 TGTAAAATTCAGCCCTTAGTCCAGATGACAGTCTTTTGTAGTGGCTCAAGTGAATGAAGC 1147
QY 1143 TGCTTACATATGGAAGGTCTCCACAACCTGGCAACCTCTACTGTGCTCTGGGTCAATTC 1202
Db 1148 TGCTTACATATGGAAGGTCTCCACAACCTGGCAACCTCTACTGTGCTCTGGGTCAATTC 1207
QY 1203 TCAGAGGTCAAGTCTGTGCTGGTGTCCATCTGACTTCAAAAGATTTGCTACCTGTTC 1262
Db 1208 TCAGAGGTCAAGTCTGTGCTGGTGTCCATCTGACTTCAAAAGATTTGCTACCTGTTC 1267
QY 1263 TGATGACAATACACTTAAATAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGGAGG 1322
Db 1268 TGATGACAATACACTTAAATAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGGAGG 1327
QY 1323 TGATAAACTTTTCCAGCGTGGGCTCTCAGAAAGAAAAAAGAGTCAAGACCTGGGCT 1382
Db 1328 TGATAAACTTTTCCAGCGTGGGCTCTCAGAAAGAAAAAAGAGTCAAGACCTGGGCT 1387
QY 1383 AGTAAACAGTAAAGAGTAGCCAGAGTACTCTGCTGCAAAAGCCCGCAGGGTAAAGTGCAATCC 1442
Db 1388 AGTAAACAGTAAAGAGTAGCCAGAGTACTCTGCTGCAAAAGCCCGCAGGGTAAAGTGCAATCC 1447
QY 1443 ATCCAAATTTCTCCCGCTCATCGCAGCTTGTGCCCAAGCTGTGTGAGAGCTTCCCTCT 1502
Db 1448 ATCCAAATTTCTCCCGCTCATCGCAGCTTGTGCCCAAGCTGTGTGAGAGCTTCCCTCT 1507
QY 1503 TCCTTCAAAATACTCTCAGTGTCTCTTATTTAAACCTCTCTGCTGCAAGGCCGCTCTCCAT 1562
Db 1508 TCCTTCAAAATACTCTCAGTGTCTCTTATTTAAACCTCTCTGCTGCAAGGCCGCTCTCCAT 1567
```

QY	1563	CAACAGAGAGGCTCTGTCTCTCTCGTCTCTCCCAAGCCACTTTCATCTTCAAGATGTC	1622
Db	1568	CAACAGAGAGGCTCTGTCTCTCTCGTCTCTCCCAAGCCACTTTCATCTTCAAGATGTC	1627
QY	1623	GATTAGAACTGGGTGAGCCGAAACACCTTCTCATCACCAACCACCTTCCACCTGCTTC	1682
Db	1628	GATTAGAACTGGGTGAGCCGAAACACCTTCTCATCACCAACCACCTTCCACCTGCTTC	1687
QY	1683	GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTCTGTGAGCCGAGATCATCCCA	1742
Db	1688	GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTCTGTGAGCCGAGATCATCCCA	1747
QY	1743	AGCAGAGCTTCTCTGAGTCTAGAAATAGATGAAGAGGAGCTAGACTCAAGCTGCT	1802
Db	1748	AGCAGAGCTTCTCTGAGTCTAGAAATAGATGAAGAGGAGCTAGACTCAAGCTGCT	1807
QY	1803	GGAGAGTGTGAAACAAAAGTGTGAAGTGTGAACCTGTGTGACTGAGCTTGTATGGCCA	1862
Db	1808	GGAGAGTGTGAAACAAAAGTGTGAAGTGTGAACCTGTGTGACTGAGCTTGTATGGCCA	1867
QY	1863	AGTTGAAATCTTCAATTTGGATCTGTGCTGCTTGTCTGTAAACGAGGAGACCTTGTATPAA	1922
Db	1868	AGTTGAAATCTTCAATTTGGATCTGTGCTGCTTGTCTGTAAACGAGGAGACCTTGTATPAA	1927
QY	1923	GGACTCTAGCTCTACCAATCAAGCAAAATTCAGAGAGCTGTACAGTATCTCAGA	1982
Db	1928	GGACTCTAGCTCTACCAATCAAGCAAAATTCAGAGAGCTGTACAGTATCTCAGA	1987
QY	1983	GCCTCGCTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGGAAGCTTACCTCTTCTT	2042
Db	1988	GCCTCGCTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGGAAGCTTACCTCTTCTT	2047
QY	2043	GAGACTTGTGTGAGAGGCTGTGAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGATPAA	2102
Db	2048	GAGACTTGTGTGAGAGGCTGTGAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGATPAA	2107
QY	2103	AAACTGGTGTGTGGCCATGGCCAGCCCAACCGGAGCTGAGATCCATCTCCACGAGTCC	2162
Db	2108	AAACTGGTGTGTGGCCATGGCCAGCCCAACCGGAGCTGAGATCCATCTCCACGAGTCC	2167
QY	2163	GTCACTCCAGACACCAATTTCCAGAGACAGAGCCGAAAGACATTCGCAAGCCCGTCCAC	2222
Db	2168	GTCACTCCAGACACCAATTTCCAGAGACAGAGCCGAAAGACATTCGCAAGCCCGTCCAC	2227
QY	2223	CATCAGCCCGAGCTCCATGAGGAAATCTGCACATATTCATATAGAAAGTCCCGAGGGA	2282
Db	2228	CATCAGCCCGAGCTCCATGAGGAAATCTGCACATATTCATATAGAAAGTCCCGAGGGA	2287
QY	2283	CTTCTGTGCTGCTGAACACTCAACAGAAATATAGATTTCTATCTGAGTGAATCTGAGC	2342
Db	2288	CTTCTGTGCTGCTGAACACTCAACAGAAATATAGATTTCTATCTGAGTGAATCTGAGC	2347
QY	2343	TTTGGTCCACTAAAACAGCTGAGCTTTGGTCCACTAAAACAAAGATGAAATAACAAG	2402
Db	2348	TTTGGTCCACTAAAACAGCTGAGCTTTGGTCCACTAAAACAAAGATGAAATAACAAG	2407
QY	2403	TGACTCTATACTCTGGTCTTTTGAAGAGCTGCCCTTTTCAATTTTATGACAAATCTTTTC	2462
Db	2408	TGACTCTATACTCTGGTCTTTTGAAGAGCTGCCCTTTTCAATTTTATGACAAATCTTTTC	2467
QY	2463	AACGCTGAAATGTACTATCTGGTCTTACTACCAATATGATATGACAGCTTCCCGAGA	2522
Db	2468	AACGCTGAAATGTACTATCTGGTCTTACTACCAATATGATATGACAGCTTCCCGAGA	2527
QY	2523	TGAATGCTGTGTTTAAATTTTCAAAAGTAAATTTGTCTACTCTGACATTTTGAATGAATAG	2582
Db	2528	TGAATGCTGTGTTTAAATTTTCAAAAGTAAATTTGTCTACTCTGACATTTTGAATGAATAG	2587
QY	2583	TCTTCACTTTTAAATTTATCTATCTTCTATAATAATGACATCCAGTTCATGAGGGA	2642
Db	2588	TCTTCACTTTTAAATTTATCTATCTTCTATAATAATGACATCCAGTTCATGAGGGA	2647
QY	2643	AAAAACAAGTTTCTTGTATTCTCTGAAACTTTTATATGCTCAGTGAAAGTATCTGCCAGCC	2702
Db	2648	AAAAACAAGTTTCTTGTATTCTCTGAAACTTTTATATGCTCAGTGAAAGTATCTGCCAGCC	2707
QY	2703	ACAGCATGAGGCTGTGAAAGCTGACTGAGAAATCTCTGCTGTAAGACCCCTGGTTCTGT	2762
Db	2708	ACAGCATGAGGCTGTGAAAGCTGACTGAGAAATCTCTGCTGTAAGACCCCTGGTTCTGT	2767
QY	2763	TCTGCTCTCCACATGATATAATTTTATTTGAAATACATAATCTTTTCACTATG	2814
Db	2768	TCTGCTCTCCACATGATATAATTTTATTTGAAATACATAATCTTTTCACTATG	2819
RESULT 4			
CQ493832			
LOCUS	CQ493832	4422 bp	DNA linear PAT 30-JAN-2004
DEFINITION	Sequence 25699 from Patent WO0160860.		
ACCESSION	CQ493832		
VERSION	CQ493832.1	GI:41459451	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Schlegel, R., Endege, W.O. and Monahan, J.E.		
TITLE	Genes differentially expressed in human prostate cancer and their use		
JOURNAL	Patent: WO 0160860-A 25699 23-AUG-2001;		
FEATURES	Millennium Predictive Medicine, Inc. (US)		
source	Location/Qualifiers		
ORIGIN	1..4422		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Query Match	99.0%;	Score 2804;	DB 2; Length 4422;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 2807;	Conservative 0;	Mismatches 5;	Indels 0; Gaps 0;
QY	3	CACGAGCGGGAGTTGGAGGCGATAACGATTTGTGTGTGAGAGCGCAACCTGGGATTTTC	62
Db	8	CAGTGGCGGGAGTTGGAGGCGATAACGATTTGTGTGTGAGAGCGCAACCTGGGATTTTC	67
QY	63	TGCTGAACTTGGAGGCAATTTCTACACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCT	122
Db	68	TGCTGAACTTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCT	127
QY	123	GATGCTCTTCAATTCGGTGTCTCCGAGCCCGAGCTTGGCGTCTGAGAAATGGATGGTC	182
Db	128	GATGCTCTTCAATTCGGTGTCTCCGAGCCCGAGCTTGGCGTCTGAGAAATGGATGGTC	187
QY	183	TTCAAAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA	242
Db	188	TTCAAAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA	247
QY	243	CACCTTTATGAGAAACAGGAGTCCGAGTTCTCTTTTGGATGTACCTTCTCTTCTGC	302
Db	248	CACCTTTATGAGAAACAGGAGTCCGAGTTCTCTTTTGGATGTACCTTCTCTTCTGC	307
QY	303	TCCCAATATGAAACATGATAGCAGTTGCCAATGAGAGGCTTTGTTCCGATTTGATAA	362
Db	308	TCCCAATATGAAACATGATAGCAGTTGCCAATGAGAGGCTTTGTTCCGATTTGATAA	367
QY	363	CACAGAAATCAAAAGTTTCAGAAAGAGTGTCTCAAGAAATGGATGGTCTCACTGGAATGC	422
Db	368	CACAGAAATCAAAAGTTTCAGAAAGAGTGTCTCAAGAAATGGATGGTCTCACTGGAATGC	427
QY	423	CGTCTTTGACCTGCGCTGGGTTCTGGTGAACTTAACTTTTACAGCAGCAGGATGATCA	482
Db	428	CGTCTTTGACCTGCGCTGGGTTCTGGTGAACTTAACTTTTACAGCAGCAGGATGATCA	487


```
|||||
2648 AAAAAAAGTTTCTTGTTATCTCTGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCACC 2707
QY
2703 ACAGCATGAGGCTCTGGAAGGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCTGT 2762
Db
2708 ACAGCATGAGGCTCTGGAAGGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCTGT 2767
QY
2763 TCTGCTCACAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db
2768 TCTGCTCACAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTATG 2819

RESULT 5
AX188263
LOCUS AX188263 4422 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 3958 from Patent WO0142467.
ACCESSION AX188263
VERSION AX188263.1 GI:15139736
KEYWORDS
SOURCE Homo sapiens (human)
% ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
AUTHORS Genes, compositions, kits, and methods for identification,
TITLE assessment, prevention, and therapy of cervical cancer
JOURNAL Patent: WO 0142467-A 3958 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source
1..4422
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.0%; Score 2804; DB 2; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACGAGCGGGAGTTGGAGCGATACGATTTGTGTGTGAGAGGCGCAACGTGCGATTTTC 62
Db 8 CAGTGGCGGAGTTGGAGCGATACGATTTGTGTGTGAGAGGCGCAACGTGCGATTTTC 67
QY 63 TGCTGAACTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 122
Db 68 TGCTGAACTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 127
QY 123 GATGCTCTTCAATTCGGTCTCGGCGAGCCCGAGCTTGGCGTCTGAGAAATGGATGTC 182
Db 128 GATGCTCTTCAATTCGGTCTCGGCGAGCCCGAGCTTGGCGTCTGAGAAATGGATGTC 187
QY 183 TTCACAAATACCCCTTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 242
Db 188 TTCACAAATACCCCTTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 247
QY 243 CACTTCTTTATGAGAAACAGGAGTCCCAAGTTCCTCTCTTTTGGATGATACCTTCTCTTCGC 302
Db 248 CACTTCTTTATGAGAAACAGGAGTCCCAAGTTCCTCTCTTTTGGATGATGATCTCTCTTCGC 307
QY 303 TCCCAATATGGAACATGTACTAGCAGTTGGCCAAATGAGAGGCTTTGTTGATTTGATATAA 362
Db 308 TCCCAATATGGAACATGTACTAGCAGTTGGCCAAATGAGAGGCTTTGTTGATTTGATATAA 367
QY 363 CACAGAAATCACAAAGTTTTCAGAAAGAGTGCTTCAAGAAATGGATGGCTCACTGGAAATGC 422
Db 368 CACAGAAATCACAAAGTTTTCAGAAAGAGTGCTTCAAGAAATGGATGGCTCACTGGAAATGC 427
QY 423 CGTCTTTGACCTGGCTGGGTTCTCGTGGTGAACCTTAAACCTTGTATACAGCAGCAGGTGATCA 482
Db 428 CGTCTTTGACCTGGCTGGGTTCTCGTGGTGAACCTTAAACCTTGTATACAGCAGCAGGTGATCA 487
```

```
QY 483 AACAGCCAAATTTTGGGACGCTAAAGAGCTGGTGAGCTGATTGGAAACATGCAAAAGGTCAATCA 542
Db 488 AACAGCCAAATTTTGGGACGCTAAAGAGCTGGTGAGCTGATTGGAAACATGCAAAAGGTCAATCA 547
QY 543 ATGCAAGCTCTCAAGTCAAGTTCGCTTTTCTAAAGTTTGGAGAAAGCTGTATTTCTGTACGGGTGG 602
Db 548 ATGCAAGCTCTCAAGTCAAGTTCGCTTTTCTAAAGTTTGGAGAAAGCTGTATTTCTGTACGGGTGG 607
QY 603 AAGAGATGGCAACATTTATGCTGCGGATACAGAGTGCAACAAAGAAAGAGTGGGTTTATAG 662
Db 608 AAGAGATGGCAACATTTATGCTGCGGATACAGAGTGCAACAAAGAAAGAGTGGGTTTATAG 667
QY 663 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTCAGACAAGCAAAACCCCTTCAAAACC 722
Db 668 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTCAGACAAGCAAAACCCCTTCAAAACC 727
QY 723 CAAGAAGAAACAGAAATTTCAAAAGAGCTTGTCTCTTCTGTGGATTTCCAGCAAAAGTGTATAC 782
Db 728 CAAGAAGAAACAGAAATTTCAAAAGAGCTTGTCTCTTCTGTGGATTTCCAGCAAAAGTGTATAC 787
QY 783 TGTGGTCTCTTTTCAAGCAGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAT 842
Db 788 TGTGGTCTCTTTTCAAGCAGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAT 847
QY 843 CAAAGTATGGGATTTTACGTAAGATTTATCTGCTTATCGACAAGAACCCATAGCATCAAA 902
Db 848 CAAAGTATGGGATTTTACGTAAGATTTATCTGCTTATCGACAAGAACCCATAGCATCAAA 907
QY 903 GTCTTTCTCTGTACCAGGTAGCAGCACTCGAAACCTTTGGATATTTCAAGTCTGATTTTGA 962
Db 908 GTCTTTCTCTGTACCAGGTAGCAGCACTCGAAACCTTTGGATATTTCAAGTCTGATTTTGA 967
QY 963 TTCCATCTGGCTCTACTTTTATTTGCTAAATTTGCAAGACGATTAACATCTACATGTTTAAAT 1022
Db 968 TTCCATCTGGCTCTACTTTTATTTGCTAAATTTGCAAGACGATTAACATCTACATGTTTAAAT 1027
QY 1023 GACTGGGTTGAAGACTTCTCCAGTGCGTATTTTCAATGGACACAGAACTCTACCTTTTA 1082
Db 1028 GACTGGGTTGAAGACTTCTCCAGTGCGCTATTTTCAATGGACACAGAACTCTACCTTTTA 1087
QY 1083 TGTAAAAATCCAGCCCTTAGTCCAGATGACCAAGTCTTGTAGTCAAGTGGCTCAAGTGAAGC 1142
Db 1088 TGTAAAAATCCAGCCCTTAGTCCAGATGACCAAGTCTTGTAGTCAAGTGAAGC 1147
QY 1143 TGCCTACATATGGAAGGTCTCCACACCTTGGCAACCTCTCTACTGCTCTCTGGGTCAATC 1202
Db 1148 TGCCTACATATGGAAGGTCTCCACACCTTGGCAACCTCTCTACTGCTCTCTGGGTCAATC 1207
QY 1203 TCAAGAGTCAAGTCTGTGTGCTGGTGCATCTGACTTTCACAAAGATTTGCTACTCTGTC 1262
Db 1208 TCAAGAGTCAAGTCTGTGTGCTGGTGCATCTGACTTTCACAAAGATTTGCTACTCTGTC 1267
QY 1263 TGATGACAATACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACGAGG 1322
Db 1268 TGATGACAATACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACGAGG 1327
QY 1323 TGATAAACTTTTCCCGGTGGGCTCTCAGAGAAAGAGTCAAGACCTTGGCT 1382
Db 1328 TGATAAACTTTTCCCGGTGGGCTCTCAGAGAAAGAGTCAAGACCTTGGCT 1387
QY 1383 AGTAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGT 1442
Db 1388 AGTAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGT 1447
QY 1443 ATCCAAATTTTCCCGGTGGGCTCTCAGAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGT 1502
Db 1448 ATCCAAATTTTCCCGGTGGGCTCTCAGAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGT 1507
QY 1503 TCCTTCAAAATCTCTAGTCTCTTATTAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGT 1562
Db 1508 TCCTTCAAAATCTCTAGTCTCTTATTAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGT 1567
QY 1563 CAACAGAAAGGCTCTGTCTCTCTCCGCTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTC 1622
```

1568 CAACGAGAGGCTCTGTCTCTCTCGTCTCTCCAGGCCACCTTCATCTTTCAAGATGTC 1627
1623 GATTAGAACTGGTGACCGCAACACCTTCCTCATCACACCCATCACTCCACCTGCTTC 1682
1628 GATTAGAACTGGTGACCGCAACACCTTCCTCATCACACCCATCACTCCACCTGCTTC 1687
1683 GGAGACCAAGATCATCTCTCCGAGAAAGCCCTTATCTGTGAGCCAGCAAGTCATCCCA 1742
1688 GGAGACCAAGATCATCTCTCCGAGAAAGCCCTTATCTGTGAGCCAGCAAGTCATCCCA 1747
1743 AGCAGAGGCTTGTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT 1802
1748 AGCAGAGGCTTGTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT 1807
1803 GGAGAGTGTGAACAAACAAAGTGTGAAGAGTGTAACTGTGACATGAGCTTGTAGGCCA 1862
1808 GGAGAGTGTGAACAAACAAAGTGTGAAGAGTGTAACTGTGACATGAGCTTGTAGGCCA 1867
1863 AGTTGAAATCTTCAATTTGGATCTGTGCTGCTTGTGCTGTAAACAGGAAGACCTTAGTAA 1922
1868 AGTTGAAATCTTCAATTTGGATCTGTGCTGCTTGTGCTGTAAACAGGAAGACCTTAGTAA 1927
1923 GGACTCTCTAGGCTCTACCAATCAAGCAAAATTTGAAGGAGCTGTGTACCAAGTATCTCAGA 1982
1928 GGACTCTCTAGGCTCTACCAATCAAGCAAAATTTGAAGGAGCTGTGTACCAAGTATCTCAGA 1987
1983 GCTCGGCTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGAACGCTACCTCTTCTCTTT 2042
1988 GCTCGGCTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGAACGCTACCTCTTCTCTTT 2047
2043 GAGACCTCTGTGAGAGGCTCTGAATGTGTGCAAGAGAGATAGTTCCCCAGAGAAATAA 2102
2048 GAGACCTCTGTGAGAGGCTCTGAATGTGTGCAAGAGAGATAGTTCCCCAGAGAAATAA 2107
2103 AAAGCTGTTGTTGGCCATCGCCCAACCGAAGGCTGAGAAATCCATCTCCACGAGTCC 2162
2108 AAAGCTGTTGTTGGCCATCGCCCAACCGAAGGCTGAGAAATCCATCTCCACGAGTCC 2167
2163 GTCATCCAGACACCAATTCAGGAGACAGAGCGGAAAGAGATTGCCAAGCCCGGTAC 2222
2168 GTCATCCAGACACCAATTCAGGAGACAGAGCGGAAAGAGATTGCCAAGCCCGGTAC 2227
2223 CATCAGCCAGCTCCATCAGAGAAATCTGCACATCTCCATAGAAAGTCCAGGAGGA 2282
2228 CATCAGCCAGCTCCATCAGAGAAATCTGCACATCTCCATAGAAAGTCCAGGAGGA 2287
2283 CTTCTGTGCTCTGAACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGC 2342
2288 CTTCTGTGCTCTGAACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGC 2347
2343 TTTGGTCCATAAACAAGCTGAGCTTTGGTCCATAAACAAGATGAAATAAACAAG 2402
2348 TTTGGTCCATAAACAAGCTGAGCTTTGGTCCATAAACAAGATGAAATAAACAAG 2407
2403 TCAGCTATAACTCTGCTCTTTAAGAAAGCTGCTTTTCATTTTGTAGCAAAATCTTTTC 2462
2408 TCAGCTATAACTCTGCTCTTTAAGAAAGCTGCTTTTCATTTTGTAGCAAAATCTTTTC 2467
2463 AACGCTGAAATGTACCTAATCTGTTTCTACTACCATTAATATATGACGCTTTCCCGAGGA 2522
2468 AACGCTGAAATGTACCTAATCTGTTTCTACTACCATTAATATATGACGCTTTCCCGAGGA 2527
2523 TGAATGCTGTGTTTAAATTTCAAAAGTAAATTTGTCACTCTAGCAATTTGAATGAATAG 2582
2528 TGAATGCTGTGTTTAAATTTCAAAAGTAAATTTGTCACTCTAGCAATTTGAATGAATAG 2587
2583 TCTTCACTTTTAAATTTCAATCTCTCTATAATAATGACATCCAGTTTCAATGAGGCA 2642
2588 TCTTCACTTTTAAATTTCAATCTCTCTATAATAATGACATCCAGTTTCAATGAGGCA 2647
2643 AAAAAACAAGTTCTTGTGTTATCTGAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702

2648 AAAAAACAAGTTCTTGTGTTATCTCGAAACTTTTCTATGTCTCAGTGGAAAGTATCTGCCAGCC 2707
2703 ACAGCATGAGGCTGTGAGGCTGACTGAGAAATCTCTGCTGAGAGCCCTGTTCTGT 2762
2708 ACAGCATGAGGCTGTGAGGCTGACTGAGAAATCTCTGCTGAGAGCCCTGTTCTGT 2767
2763 TCTGCTCTCAACATGATATAATTTATTTTCAATAATCATATCTTTTCACTATG 2814
2768 TCTGCTCTCAACATGATATAATTTATTTTGAATACATAATCTTTTCACTATG 2819

RESULT 6
AK000742
LOCUS Homo sapiens cDNA FLJ20735 fis, clone HEP08479. linear PRI 13-SEP-2003
DEFINITION AK000742
ACCESSION AK000742
VERSION AK000742.1 GI:7021020
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Kawakami, T., Nouchi, S., Itoh, T., Shigeta, K., Senba, T.,
Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Ota, T.,
Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
2 (bases 1 to 3008)
Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-ku, Tokyo 108-8639, Japan (E-mail: flicdna@ims.u-tokyo.ac.jp,
Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
Location/Qualifiers
1. 3008
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP08479"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/note="Cloning vector pME18SFL3"
157..2349
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91355.1"
/trnslat="MLFNSVLRLQGLVLRNGSSQYPLQSLITGTCQSGNDEHTSYG
ETGVPVPPFGCTSSAANMEHLAVANEGBVRLYNTEPQSPKXKPKEMWAHNAVF
DLAWPGLKLTWAGDOTAKFWDVKAGELIGTCKGHQCSLKSVAFKPEKAVFTGG
RDGNIMWDTKCNKDGFRQVNOISGAHNTSDKQTPSKPKKQNSKSLAPSVDFQQS
VTYVLFODENTLVAGAVDGIKVMDLRKNYTAHQEPYASKSLYPCSSSTRKLGYS
LILDSSTLSTFANCTDDNIYFMNTGLKTSFPAIFNGHNSFTFYKSSLSDDQPLVS
GSSDEAAVIMKVSTPWPPTVLLGHOSBTVSCWPCSDFTKIATCDDNTLKIWLNR
GLEKPGDKLSTVWASOKKESRPGIIVTSSQSTPAKAPRVKCNPSNPSAAAC
APSCAGDULPSPNTPTFSIKTSPAKAPINRRGSSVSVPKPPSSFKMSIRNWTRT
PSSPPPTTPPASETKIMSPKRALIPVQKSSQAEACSESRRNRVRRRLSDCLSVKQK
PSSKNCVTELDGVENHLDLCLLQAGNEDLSKDSLGTPTKSIKIEGATGTSIEPSP
ISPVSBSGCTPLPLRPGSGSMVGKNSPENKWLMLAMAAKRAENPSRSPSS
QTPNSRSGKTLPSVTLTPTSSMRKICTYFHRKQBEDFCGPEHSTL"

ORIGIN																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					</
--------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----


```
QY 2103 AACTCGTGTGTTGGCCATGCGCAGCAAAACGGAAGGCTGAGATCCATCTCCACGAAGTCC 2162
Db 2136 AACTCGTGTGTTGGCCATGCGCAGCAAAACGGAAGGCTGAGATCCATCTCCACGAAGTCC 2195
QY 2163 GTCATCCGACACCCCAATTCGAGGAGACAGAGCGGAAGACATTCGCCAAGCCCGGTAC 2222
Db 2196 GTCATCCGACACCCCAATTCGAGGAGACAGAGCGGAAGACATTCGCCAAGCCCGGTAC 2255
QY 2223 CATCAGCCCGAGCTCCATGAGGAAATCTGCACATCTTCATAGAAAGTCCAGGAGGA 2282
Db 2256 CATCAGCCCGAGCTCCATGAGGAAATCTGCACATCTTCATAGAAAGTCCAGGAGGA 2315
QY 2283 CTTCTGTGCTGCTGAACACTCAACAGATTAATAGATTCTAATCTGAGTGAGTTACTGAGC 2342
Db 2316 CTTCTGTGCTGCTGAACACTCAACAGATTAATAGATTCTAATCTGAGTGAGTTACTGAGC 2375
QY 2343 TTGTGTCCTAAACAAAGCTGAGCTTTGGTTCCTACTACATAATGTATATGACGCTTCCCGAGGA 2402
Db 2376 TTGTGTCCTAAACAAAGCTGAGCTTTGGTTCCTACTACATAATGTATATGACGCTTCCCGAGGA 2435
QY 2403 TGACTCTATAACTCTGCTTTTAAAGAAAGCTGCTTTTCAATTTTACACAAATCTTTTC 2462
Db 2436 TGACTCTATAACTCTGCTTTTAAAGAAAGCTGCTTTTCAATTTTACACAAATCTTTTC 2495
QY 2463 AAGCTGAAATGTACCTAATCTGTTTCTACTACATAATGTATATGACGCTTCCCGAGGA 2522
Db 2496 AAGCTGAAATGTACCTAATCTGTTTCTACTACATAATGTATATGACGCTTCCCGAGGA 2555
QY 2523 TGAATGCTGTGTTTAAATTTCAATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAG 2582
Db 2556 TGAATGCTGTGTTTAAATTTCAATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAG 2615
QY 2583 TCTTCACTTTTAAATTTTAAATTTTCTCTAATAATGACATCCAGTTTCAATGAGGCA 2642
Db 2616 TCTTCACTTTTAAATTTTAAATTTTCTCTAATAATGACATCCAGTTTCAATGAGGCA 2675
QY 2643 AAAAAACAAGTTCTGTTTATCTCTGAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702
Db 2676 AAAAAACAAGTTCTGTTTATCTCTGAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2735
QY 2703 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCTCTGCTGAGAACCCCTGGTTCTGT 2762
Db 2736 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCTCTGCTGAGAACCCCTGGTTCTGT 2795
QY 2763 TCTGCTCCCAACATGATATTTTATTTTCAATACATATCTTTTCACTATG 2814
Db 2796 TCTGCTCCCAACATGATATTTTATTTTGAATAACATATCTTTTCACTATG 2847

RESULT 7
LOCUS BC033540
DEFINITION Homo sapiens denticleless homolog (Drosophila), mRNA (cdna clone
ACCESSION BC033540
VERSION MG:42657 IMAGE:4826434), complete cds.
KEYWORDS BC033540.1 GI:23272356
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2969)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
```

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketterman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shvetchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Mammalian Gene Collection Program Team

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2969)

NIH MGC Project

Direct Submission

Submitted (26-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 71 Row: 0 Column: 18.

FEATURES

source

1..2969

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:42657 IMAGE:4826434"

/tissue_type="Testis"

/clone_lib="NIH-MGC_97"

/lab_host="DH10B"

/note="Vector: pBluescriptR"

1..2969

/gene="DTL"

/note="synonyms: L2DTL, RAMP"

/db_xref="GeneID:51514"

121..2313

/gene="DTL"

/codon_start=1

/product="denticleless homolog"

/protein_id="AAH3540.1"

/db_xref="GI:23272357"

/db_xref="GeneID:51514"

gene

CDS

ORIGIN

Query Match									
Best Local Similarity 99.0%; Score 2802.2; DB 5; Length 2969;									
Matches 2804; Conservative 0; Mismatches 3; Indels 0; Gaps 0;									
Qy	8	GGGGAGTTGGAGGCGATACGATTGTGTTCTGTGAGAGCGCAACGTGCGATTTCGTGTG	67						
Db	5	GGGGAGTTGGAGGCGATACGATTGTGTTGTGTGAGAGCGGAGCTGGATTTCGTGTG	64						
Qy	68	AATCTTGGAGGCAATTTCTACGACTTTTCTCTACGCTGAGGCTTTTCTCCGACCCCTGATGC	127						
Db	65	AATCTTGGAGGCAATTTCTACGACTTTTCTCTACGCTGAGGCTTTTCTCCGACCCCTGATGC	124						
Qy	128	TCTTCAATTCGGTGTCTCCGACGCCAGCTTGGGCTCTGAGAAATGATGGTCTTTCAC	187						
Db	125	TCTTCAATTCGGTGTCTCCGACGCCAGCTTGGGCTCTGAGAAATGATGGTCTTTCAC	184						
Qy	188	AATACCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAATGATGAACACATT	247						
Db	185	AATACCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAATGATGAACACATT	244						
Qy	248	CTTATGGAGAAACAGAGTCCAGTTCTCTCTTTTGGATGTACTCTCTCTGCTCCCA	307						
Db	245	CTTATGGAGAAACAGAGTCCAGTTCTCTCTTTTGGATGTACTCTCTCTGCTCCCA	304						
Qy	308	ATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTCGATGTATACACAG	367						
Db	305	ATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTCGATGTATACACAG	364						
Qy	368	AATCAAAAGTTTCAGAAAGAGTGCTTCAAGAATGATGGCTCACTGGAATGCCGTCT	427						
Db	365	AATCAAAAGTTTCAGAAAGAGTGCTTCAAGAATGATGGCTCACTGGAATGCCGTCT	424						
Qy	428	TTGACCTGCGCTGGTTCCTGGTGAACCTTAACTTTTACAGCAGCAGGTGATCAACAG	487						
Db	425	TTGACCTGCGCTGGTTCCTGGTGAACCTTAACTTTTACAGCAGCAGGTGATCAACAG	484						
Qy	488	CCAAATTTGGGACGTAAAGCTGTGAGCTGATTTGGAACATGCAAGCTCATCATGCA	547						
Db	485	CCAAATTTGGGACGTAAAGCTGTGAGCTGATTTGGAACATGCAAGCTCATCATGCA	544						
Qy	548	GCCTCAAGTCAGTGCCTTTCTTAAGTTTGAGAAAGCTGATTTCTGTCGGGTGGAAG	607						
Db	545	GCCTCAAGTCAGTGCCTTTCTTAAGTTTGAGAAAGCTGATTTCTGTCGGGTGGAAG	604						
Qy	608	ATGGCAACATTATGGTCTGGGATACCAGGTGCAACAAAAAGATGGTTTTATAGGCAAG	667						
Db	605	ATGGCAACATTATGGTCTGGGATACCAGGTGCAACAAAAAGATGGTTTTATAGGCAAG	664						
Qy	668	TGAATCAAAATCAGTGGAGCTCAATACCTCAGACAGCAAAACCCCTTCAAAACCAAGA	727						
Db	665	TGAATCAAAATCAGTGGAGCTCAATACCTCAGACAGCAAAACCCCTTCAAAACCAAGA	724						
Qy	728	AGAAACAGAAATCAAAAGCACTTGCTCTCTGTGGATTTCCAGCAAGTGTACTGTGG	787						
Db	725	AGAAACAGAAATCAAAAGCACTTGCTCTCTGTGGATTTCCAGCAAGTGTACTGTGG	784						
Qy	788	TCTCTCTTCAAGACAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATTAATCAAG	847						
Db	785	TCTCTCTTCAAGACAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATTAATCAAG	844						
Qy	848	TATGGGATTTAGTAAGATTTACTGTCTTATCGACAGAACCCATAGCATCCAAGTCTT	907						
Db	845	TATGGGATTTAGTAAGATTTACTGTCTTATCGACAGAACCCATAGCATCCAAGTCTT	904						
Qy	908	TCTGTACCCAGGTAGCAGCTCGAAACTTGGATTTCAAGTCTGATTTTGGATTTCA	967						
Db	905	TCTGTACCCAGGTAGCAGCTCGAAACTTGGATTTCAAGTCTGATTTTGGATTTCA	964						
Qy	968	CTGGCTCTACTTTTATTTGCTAATTCGACAGCGATAACATCTACATGTTTAAATGACTG	1027						
Db	965	CTGGCTCTACTTTTATTTGCTAATTCGACAGCGATAACATCTACATGTTTAAATGACTG	1024						

Qy	1028	GTTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTATGTAA	1087
Db	1025	GTTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTATGTAA	1084
Qy	1088	AATCCAGCCTTATGTCAGATGACACAGTTTATTTAGTCAGTGGCTCAAGTGATGAAGCTGCCT	1147
Db	1085	AATCCAGCCTTATGTCAGATGACACAGTTTATTTAGTCAGTGGCTCAAGTGATGAAGCTGCCT	1144
Qy	1148	ACATATGGAAGGTCTCACACCTTGGCAACCTCTCTACTGTGCTCTCTGGGTCAATCTCAAG	1207
Db	1145	ACATATGGAAGGTCTCACACCTTGGCAACCTCTCTACTGTGCTCTCTGGGTCAATCTCAAG	1204
Qy	1208	AGGTACGTCTGTGCTGTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTCTGATG	1267
Db	1205	AGGTACGTCTGTGCTGTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTCTGATG	1264
Qy	1268	ACAATACATAAAAAATCTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACACAGGAGGTGATA	1327
Db	1265	ACAATACATAAAAAATCTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACACAGGAGGTGATA	1324
Qy	1328	AACCTTCCA CGTGGGTGGGCTCTCTCAGAAAGAAAAAGAGTCAAGACCTGGCCCTAGTAA	1387
Db	1325	AACCTTCCA CGTGGGTGGGCTCTCTCAGAAAGAAAAAGAGTCAAGACCTGGCCCTAGTAA	1384
Qy	1388	CAGTAAAGAGTAGCAGAGTACTCTGCCAAAGCCCCCAGGGTAAAGTGCAATCCATCCA	1447
Db	1385	CAGTAAAGAGTAGCAGAGTACTCTGCCAAAGCCCCCAGGGTAAAGTGCAATCCATCCA	1444
Qy	1448	ATTCTTCCCCTCATCCGCGAGCTTGTGCCCAAGCTGTGTGGAGACCTCCCTCTTCCTT	1507
Db	1445	ATTCTTCCCCTCATCCGCGAGCTTGTGCCCAAGCTGTGTGGAGACCTCCCTCTTCCTT	1504
Qy	1508	CAAAATCTCTACGTCTCTATTTAAAACCTCTCTCTGCCAAGGCCGGTCTCCCATCAACA	1567
Db	1505	CAAAATCTCTACGTCTCTATTTAAAACCTCTCTCTGCCAAGGCCGGTCTCCCATCAACA	1564
Qy	1568	GAAGAGCTCTGTCTCTCCGCTCTCTCCCAAGCCACCTTCACTTTCAAGATGTCGATTA	1627
Db	1565	GAAGAGCTCTGTCTCTCCGCTCTCTCCCAAGCCACCTTCACTTTCAAGATGTCGATTA	1624
Qy	1628	GAATACTGGGTGACCGGAACCTTCTCTCATCAACCAACCATCACTCCACCTCTTCGGAGA	1687
Db	1625	GAATACTGGGTGACCGGAACCTTCTCTCATCAACCAACCATCACTCCACCTCTTCGGAGA	1684
Qy	1688	CCAAGATCATGTCTCCGAGAAAAAGCCCTTATCTCTGTGAGCCAGAGTCATCCCAAGCAG	1747
Db	1685	CCAAGATCATGTCTCCGAGAAAAAGCCCTTATCTCTGTGAGCCAGAGTCATCCCAAGCAG	1744
Qy	1748	AGGCTTGTCTGAGTCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCTGGAGA	1807
Db	1745	AGGCTTGTCTGAGTCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCTGGAGA	1804
Qy	1808	GTGTGAAAACAAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGCCAAGTTG	1867
Db	1805	GTGTGAAAACAAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGCCAAGTTG	1864
Qy	1868	AAAATCTTTCATTTGGATCTGTGCTTSCCTTGTGTTAACAGAGAACCTTATAGTAGGACT	1927
Db	1865	AAAATCTTTCATTTGGATCTGTGCTTSCCTTGTGTTAACAGAGAACCTTATAGTAGGACT	1924
Qy	1928	CTCTAGGTCCTACCAAAATCAAGCAAAATTTGAAGAGCTGGTACCAGATATCTCAGAGCCCTC	1987
Db	1925	CTCTAGGTCCTACCAAAATCAAGCAAAATTTGAAGAGCTGGTACCAGATATCTCAGAGCCCTC	1984
Qy	1988	CGTCTCTATCAGTCCGCTATGCTTTCAGAAAGCTGTGGAAACGCTACCTCTTCTTTGAGAC	2047
Db	1985	CGTCTCTATCAGTCCGCTATGCTTTCAGAAAGCTGTGGAAACGCTACCTCTTCTTTGAGAC	2044
Qy	2048	CTTGTGGAGAGGGTCTGAAAATGGTAGGCAAGAGATAGTTTCCCAGAGATATAAAACT	2107
Db	2045	CTTGTGGAGAGGGTCTGAAAATGGTAGGCAAGAGATAGTTTCCCAGAGATATAAAACT	2104
Qy	2108	GGTTGTTGGCCATGSGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCCGTGCTAT	2167

Db 2105 GGTGTTGGCCATGGCAGCAACGAGGCTGAGAATCCATCTCCACGAAAGTCCGTCAT 2164
Qy 2168 CCAGACACCCAAATCCAGAGACAGCGGGAAGACATTCGACGCGGCTCACCATCA 2227
Db 2165 CCAGACACCCAAATCCAGAGACAGCGGGAAGACATTCGACGCGGCTCACCATCA 2224
Qy 2228 CGCCAGCTCCATAGAGAAATCTGCACATATCTCCATAGAAAGTCCAGGAGGACTTCT 2287
Db 2225 CGCCAGCTCCATAGAGAAATCTGCACATATCTCCATAGAAAGTCCAGGAGGACTTCT 2284
Qy 2288 GTGGTCTCAACACCTCAACAGAAATATAGATCTTAATCTGAGTGATTAATCTGAGCTTTGG 2347
Db 2285 GTGGTCTCAACACCTCAACAGAGTTATAGATCTTAATCTGAGTGATTAATCTGAGCTTTGG 2344
Qy 2348 TCCACTAACAACAGCTGAGCTTTGGTCCACTAACAACAGATGAATAACAAGAGTACT 2407
Db 2345 TCCACTAACAACAGCTGAGCTTTGGTCCACTAACAACAGATGAATAACAAGAGTACT 2404
Qy 2408 CTATAACTCTGGTCTTTAAGAAAGCTGCTTTTCAATTTTATAGCAAAATCTTTTCAACGC 2467
Db 2405 CTATAACTCTGGTCTTTAAGAAAGCTGCTTTTCAATTTTATAGCAAAATCTTTTCAACGC 2464
Qy 2468 TGAATGTACCTAATCTGTTCTACTA CCAATATGTATGACGCTTCCGAGGATGAAT 2527
Db 2465 TGAATGTACCTAATCTGTTCTACTA CCAATATGTATGACGCTTCCGAGGATGAAT 2524
Qy 2528 GCTGTGTTTAAATTTTCAATAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTC 2587
Db 2525 GCTGTGTTTAAATTTTCAATAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTC 2584
Qy 2588 ACTTTTAAATTTTCAATCTCTCTATATAATGACATCCAGTTTCATGAGGCAAAAAA 2647
Db 2585 ACTTTTAAATTTTCAATCTCTCTATATAATGACATCCAGTTTCATGAGGCAAAAAA 2644
Qy 2648 CAAGTTCTTGTATCTCTGAAACTTTCTATGCTCAGTGGAAGATATCTGCCAGCCACAGC 2707
Db 2645 CAAGTTCTTGTATCTCTGAAACTTTCTATGCTCAGTGGAAGATATCTGCCAGCCACAGC 2704
Qy 2708 ATGAGGCTGTGAAGCTGACTGAGAAATCTCTGCTGGAAGCCCTGTTCTGTTCTGCTG 2767
Db 2705 ATGAGGCTGTGAAGCTGACTGAGAAATCTCTGCTGGAAGCCCTGTTCTGTTCTGCTG 2764
Qy 2768 CTCCAACTGTATATTTTATTTTGAATACATAATCTTTTCACTATG 2814
Db 2765 CTCCAACTGTATATTTTATTTTGAATACATAATCTTTTCACTATG 2811

RESULT 8
AX780410
LOCUS AX780410 4450 bp DNA linear PAT 14-JUL-2003
DEFINITION Sequence 2567 from Patent WO03039443.
ACCESSION AX780410
VERSION AX780410.1 GI:32697404
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
Dugas, M., Eils, R., Brors, B. and Mergenthaler, S.
TITLE Novel genetic markers for leukemias
JOURNAL Patent: WO 03039443-A 2567 15-MAY-2003;
Deutsches Krebsforschungszentrum (DE);
Ludwig-Maximilian-Universitaet Muenchen (DE);
PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)
FEATURES
source
1..4450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.0%; Score 2802; DB 2; Length 4450;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2805; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 3 CACGAGCGGGAGTTGGAGGCGATACGATTTTGTGTGTGAGAGGCGCAACGTCGATTTTC 62
Db 36 CAGTGGCGGGAGTTGGAGGCGATACGATTTTGTGTGTGAGAGGCGCAACGTCGATTTTC 95
Qy 63 TGTGAATCTGGAGGCAATTTCTA C GACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCTT 122
Db 96 TGTGAATCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCTT 155
Qy 123 GATGCTCTTCAATTCGCTGCTCCGCCAGCCAGCTTCGCGTCTGAGAAATGGAATGGTC 182
Db 156 GATGCTCTTCAATTCGCTGCTCCGCCAGCCAGCTTCGCGTCTGAGAAATGGAATGGTC 215
Qy 183 TTCAATATACCTCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 242
Db 216 TTCAATATACCTCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 275
Qy 243 CACTTCTTATGAGAAACAGGAGTCCCAAGTTCCTCTCTTTTGGATGTACTCTTCTCTTCG 302
Db 276 CACTTCTTATGAGAAACAGGAGTCCCAAGTTCCTCTCTTTTGGATGTACTCTTCTCTTCG 335
Qy 303 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTCCGATTTGATAA 362
Db 336 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTCCGATTTGATAA 395
Qy 363 CACAGAAATCACAAGTTTTCAGAAAGAAAGTGTCTTCAAGAAATGGATGGTCTCACTGGAATGC 422
Db 396 CACAGAAATCACAAGTTTTCAGAAAGAAAGTGTCTTCAAGAAATGGATGGTCTCACTGGAATGC 455
Qy 423 CGTCTTCACTGCGCTGGTTCCTGGTGAACCTTAAACTTGTGTACAGCAGCAGGTGATCA 482
Db 456 CGTCTTCACTGCGCTGGTTCCTGGTGAACCTTAAACTTGTGTACAGCAGCAGGTGATCA 515
Qy 483 AACAGCCAAATTTTGGGACGTAAAGCTGTGAGCTGATTCGAAATCACAAGTGCATCA 542
Db 516 AACAGCCAAATTTTGGGACGTAAAGCTGTGAGCTGATTCGAAATCACAAGTGCATCA 575
Qy 543 ATGAGCCCTCAAGTCAGTTGCTTTTCTAAGTTTGGAGAAAGCTGTATTCGTAACGAGTGG 602
Db 576 ATGAGCCCTCAAGTCAGTTGCTTTTCTAAGTTTGGAGAAAGCTGTATTCGTAACGAGTGG 635
Qy 603 AAGAGATGGCAACATTTATGCTCTGGGATACAGGTGCAACAAAAAAGATGGGTTTATAG 662
Db 636 AAGAGATGGCAACATTTATGCTCTGGGATACAGGTGCAACAAAAAAGATGGGTTTATAG 695
Qy 663 GCAAGTGAATCAATCAGTGGAGCTCAATACCTCAGACAGCAAGCAACCCCTTCAAAACC 722
Db 696 GCAAGTGAATCAATCAGTGGAGCTCAATACCTCAGACAGCAAGCAACCCCTTCAAAACC 755
Qy 723 CAAGAAGAAACAGAAATTCAAAAGGACTTGTCTCTCTGAGATTTCCAGCAAGTGTTCAC 782
Db 756 CAAGAAGAAACAGAAATTCAAAAGGACTTGTCTCTCTGAGATTTCCAGCAAGTGTTCAC 815
Qy 783 TGTGCTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAT 842
Db 816 TGTGCTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAT 875
Qy 843 CAAGATATGGGATTTTACGTAGAAATTTATCTGCTTATCGACAGNAGCCATAGCATCAA 902
Db 876 CAAGATATGGGATTTTACGTAGAAATTTATCTGCTTATCGACAGNAGCCATAGCATCAA 935
Qy 903 GTCTTTCTCTGACCCAGCTAGCAGCCTCGAAATCTTGGATATTTCAAGTCTGATTTTGA 962
Db 936 GTCTTTCTCTGACCCAGCTAGCAGCCTCGAAATCTTGGATATTTCAAGTCTGATTTTGA 995
Qy 963 TTCCACTGGCTCTACTTTTATTTGCTAAATTTGACAGCAAGCAATCACTCATGTTTAAAT 1022
Db 996 TTCCACTGGCTCTACTTTTATTTGCTAAATTTGACAGCAAGCAATCACTCATGTTTAAAT 1055

Qy	1023	GACTGGGTTGAGAGCTTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTA	1082
Db	1056	GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTA	1115
Qy	1083	TGTAANAATCCAGCCTTAGTCCAGATAGCAGTTTTTTTAGTCAGTGGCTCAAGTGAAGC	1142
Db	1116	TGTAANAATCCAGCCTTAGTCCAGATAGCAGTTTTTTTAGTCAGTGGCTCAAGTGAAGC	1175
Qy	1143	TGCCTACATATGGAAGGTCTCCACACCTGGCAACCTCCTACTGTGTCTCCTGGGTCAATTC	1202
Db	1176	TGCCTACATATGGAAGGTCTCCACACCTGGCAACCTCCTACTGTGTCTCCTGGGTCAATTC	1235
Qy	1203	TCAAGAGGTCACTGTCTGTGTCTGGTGTCCATCTGAATTTCACAAAGATTTGCTACCTGTTC	1262
Db	1236	TCAAGAGGTCACTGTCTGTGTCTGGTGTCCATCTGAATTTCACAAAGATTTGCTACCTGTTC	1295
Qy	1263	TGATGCAATACACTTAANAATCTGGCCCTTGAATAGAGGCTTAGAGAGAGAAACGAGGAGG	1322
Db	1296	TGATGCAATACACTTAANAATCTGGCCCTTGAATAGAGGCTTAGAGAGAGAAACGAGGAGG	1355
Qy	1323	TGATAAACTTTCCACGGTGGGTCTGGGCTCTTCAGAGAAGAAAAAGAGTCAAGACTCGGCCT	1382
Db	1356	TGATAAACTTTCCACGGTGGGTCTGGGCTCTTCAGAGAAGAAAAAGAGTCAAGACTCGGCCT	1415
Qy	1383	AGTAACAGTAAACGATAGCCAGATGTACTCTGTGCCAAAGGCCCCAGGGTAAAGTGCATATCC	1442
Db	1416	AGTAACAGTAAACGATAGCCAGATGTACTCTGTGCCAAAGGCCCCAGGGTAAAGTGCATATCC	1475
Qy	1443	ATCCAAATTTCTCCCGTCAATCCGAGCTTGTCGCCCAAGCTGTGTCTGGAGACCTCCCTCT	1502
Db	1476	ATCCAAATTTCTCCCGTCAATCCGAGCTTGTCGCCCAAGCTGTGTCTGGAGACCTCCCTCT	1535
Qy	1503	TCCTTTCAAAATCTCCTAGTTTCTCTATTAAACCTCTCTGCCAAGGCCCGGTCTCCCAT	1562
Db	1536	TCCTTTCAAAATCTCCTAGTTTCTCTATTAAACCTCTCTGCCAAGGCCCGGTCTCCCAT	1595
Qy	1563	CAACAGAGAGGCTCTGTCTCCTCCGTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTC	1622
Db	1596	CAACAGAGAGGCTCTGTCTCCTCCGTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTC	1655
Qy	1623	GATTAGAAAACCTGGGTGACCCGAAACACTTCTCTATCAACCCCACTCACTCCCTGCTTC	1682
Db	1656	GATTAGAAAACCTGGGTGACCCGAAACACTTCTCTATCAACCCCACTCACTCCCTGCTTC	1715
Qy	1683	GGAGACCAAGATCATGTCTCCGAGAAAAGCCCTTATTCTGTGAGCCAGAGTCAATCCCA	1742
Db	1716	GGAGACCAAGATCATGTCTCCGAGAAAAGCCCTTATTCTGTGAGCCAGAGTCAATCCCA	1775
Qy	1743	AGCAGAGGCTTGCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT	1802
Db	1776	AGCAGAGGCTTGCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT	1835
Qy	1803	GGAGAGTGTGAAAACAAAAGCTGTGTGAAGAGTGTGTAACTGTGTGACTCAGCTTTGATGCCA	1862
Db	1836	GGAGAGTGTGAAAACAAAAGCTGTGTGAAGAGTGTGTAACTGTGTGACTCAGCTTTGATGCCA	1895
Qy	1863	AGTTGAAAATCTTCAATTTGGAATCTGTGCTGCTTGTGGTAAACAGGAAGACCTTAGTAA	1922
Db	1896	AGTTGAAAATCTTCAATTTGGAATCTGTGCTGCTTGTGGTAAACAGGAAGACCTTAGTAA	1955
Qy	1923	GGACTCTTAGGCTCCTACCAATCAAGCAAAATTTGAAGGAGCTGGTACCGATCTCAGA	1982
Db	1956	GGACTCTTAGGCTCCTACCAATCAAGCAAAATTTGAAGGAGCTGGTACCGATCTCAGA	2015
Qy	1983	GCCTCCGCTCCTATCAGTCCGATGCTTTCAGAAAAGCTGTGTGAAAAGCTTCTCCTTTT	2042
Db	2016	GCCTCCGCTCCTATCAGTCCGATGCTTTCAGAAAAGCTGTGTGAAAAGCTTCTCCTTTT	2075
Qy	2043	GAGACCTTTGTGGAGAGGGTCTGAAATGGTATGGCAAAAGAGAAATAGTTCCTCCAGAAATA	2102
Db	2076	GAGACCTTTGTGGAGAGGGTCTGAAATGGTATGGCAAAAGAGAAATAGTTCCTCCAGAAATA	2135

QY	2103	AAACTGGTTGTTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCC	2162
Db	2136	AAACTGGTTGTTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCC	2195
QY	2163	GTCAATCCAGACACCCAAATTCGAGGACAGAGCGGAAAGACATGTCGAAGCCCGGTCCAC	2222
Db	2196	GTCATCCAGACACCCAAATTCGAGGACAGAGCGGAAAGACATGTCGAAGCCCGGTCCAC	2255
QY	2223	CATCACGCCAGCTCCATGAGGAAATCTGCACATACTTCCATAGAAAGTCCCAGGAGGA	2282
Db	2256	CATCACGCCAGCTCCATGAGGAAATCTGCACATACTTCCATAGAAAGTCCCAGGAGGA	2315
QY	2283	CTTCTGTGTCCTCGAAACACTCAACGAAATATATAGATTTCTAATCTGTAGTGTACTCTGAGC	2342
Db	2316	CTTCTGTGTCCTCGAAACACTCAACGAAATATATAGATTTCTAATCTGTAGTGTACTCTGAGC	2375
QY	2343	TTTGGTCCACTAAAAAAGCTGTGCTTTGGTCCACTAAAAAAGATCAAAATATACAAGAG	2402
Db	2376	TTTGGTCCACTAAAAAAGCTGTGCTTTGGTCCACTAAAAAAGATCAAAATATACAAGAG	2435
QY	2403	TGACTCTATAACTCTGTGCTTTTAAGAAAGCTGCTTTTTCATTTTTTAGACAAAATCTTTTC	2462
Db	2436	TGACTCTATAACTCTGTGCTTTTAAGAAAGCTGCTTTTTCATTTTTTAGACAAAATCTTTTC	2495
QY	2463	AACCTGAAATGTACCTAATCTGCTCTACTACCAATATGTATATGATGAGCTTCCCGAGGA	2522
Db	2496	AACCTGAAATGTACCTAATCTGCTCTACTACCAATATGTATATGATGAGCTTCCCGAGGA	2555
QY	2523	TGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGTCACTTAGCATTTTGAATGAATAG	2582
Db	2556	TGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGTCACTTAGCATTTTGAATGAATAG	2615
QY	2583	TCTTCACTTTTTTAAATTTATTCACCTCTCTATATAATATGACATCCCAGTTTCATGGAGCA	2642
Db	2616	TCTTCACTTTTTTAAATTTATTCACCTCTCTATATAATATGACATCCCAGTTTCATGGAGCA	2675
QY	2643	AAACAACAGTTTCTGTGTTATCTCTGAAACTTTCTATGCTCAGTGAAGAAAGTATCTGCCAGCC	2702
Db	2676	AAACAACAGTTTCTGTGTTATCTCTGAAACTTTCTATGCTCAGTGAAGAAAGTATCTGCCAGCC	2735
QY	2703	ACAGCATGAGGCTGTGAAGGCTGATGAGAAATCTCTGCTGAAGACCCCTGGTTCTGT	2762
Db	2736	ACAGCATGAGGCTGTGAAGGCTGATGAGAAATCTCTGCTGAAGACCCCTGGTTCTGT	2795
QY	2763	TCTGCCTCCAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTA	2812
Db	2796	TCTGCCTCCAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTA	2845
RESULT 9			
AX928533			
LOCUS	AX928533	4221 bp	DNA linear PAT 19-DEC-2003
DEFINITION	Sequence 85 from Patent EP1355150.		
ACCESSION	AX928533		
VERSION	AX928533.1	GI:40246344	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Wang, Y.		
TITLE	Panel of nucleic acid sequences for cancer diagnosis		
JOURNAL	Patent: EP 1355150-A 85 22-OCT-2003;		
FEATURES	Ortho-Clinical Diagnostics, Inc. (US)		
source	Location/Qualifiers		
	1. 4221		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			

Query Match									
Best Local Similarity 98.5%; Score 2789.8; DB 2; Length 4221;									
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	22	CGATACGATTTGTTGTGAGAGGCGCAACGTGGATTTCTGCTGAACCTTGGAGGCATT	81						
DB	1	CGATAACGATTTGTTGTGAGAGGCGCAAGCTGGCAATTTCTGCTGAACCTTGGAGGCATT	60						
QY	82	TCCTAGACCTTTCTCTCAGCTGAGGCTTTCTCCGACCCCTGATGCTCTCAATTCGGTG	141						
DB	61	TCCTAGACCTTTCTCTCAGCTGAGGCTTTCTCCGACCCCTGATGCTCTCAATTCGGTG	120						
QY	142	CTCCGCGAGCCCGAGCTTGGCGTCTCTGAGAAATGGATGGTCTTCAATACCCCTCTTCAA	201						
DB	121	CTCCGCGAGCCCGAGCTTGGCGTCTCTGAGAAATGGATGGTCTTCAATACCCCTCTTCAA	180						
QY	202	TCCCTTCTGACTGGTTATCAGTGCAGTGGAATGATGAACAACACTTCTTATGGAGAAACA	261						
DB	181	TCCCTTCTGACTGGTTATCAGTGCAGTGGAATGATGAACAACACTTCTTATGGAGAAACA	240						
QY	262	GGAGTCCCAGTTCCTCTTTTGGATGTACTTCTCTCTGCTCCCAATATGAACATGTA	321						
DB	241	GGAGTCCCAGTTCCTCTTTTGGATGTACTTCTCTCTGCTCCCAATATGAACATGTA	300						
QY	322	CTAGCAGTTGCCAATGAAGAAGCTTTTGTGANTGTATAACACAGAAATCAAAAAGTTTC	381						
DB	301	CTAGCAGTTGCCAATGAAGAAGCTTTTGTGANTGTATAACACAGAAATCAAAAAGTTTC	360						
QY	382	AGAAAGAGTGTCTCAAAGATGGATGGCTCACTGGAATGCCGTCTTGAACCTGGCTGG	441						
DB	361	AGAAAGAGTGTCTCAAAGATGGATGGCTCACTGGAATGCCGTCTTGAACCTGGCTGG	420						
QY	442	GTCTCTGTGAACTTAACTCTTACACAGCAGAGTGATCAACAGCCCAAAATTTTGGGAC	501						
DB	421	GTCTCTGTGAACTTAACTCTTACACAGCAGAGTGATCAACAGCCCAAAATTTTGGGAC	480						
QY	502	GTAAAGCTGGTGAGCTGATGGAACATGCAAGAGTCAATGACAGCCCTCAAGTCAGTT	561						
DB	481	GTAAAGCTGGTGAGCTGATGGAACATGCAAGAGTCAATGACAGCCCTCAAGTCAGTT	540						
QY	562	GCCTTTTCTAACTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGATGGCAACATTATG	621						
DB	541	GCCTTTTCTAACTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGATGGCAACATTATG	600						
QY	622	GTCTGGGATACAGGTGCAACAAAAGATGGTGTATAGGCAAGTGAATCAATCAGT	681						
DB	601	GTCTGGGATACAGGTGCAACAAAAGATGGTGTATAGGCAAGTGAATCAATCAGT	660						
QY	682	GGAGCTCAAAATACCTCAGACAAGCAACCCCTTCAAAACCCCAAGAAAGACAGAAATCA	741						
DB	661	GGAGCTCAAAATACCTCAGACAAGCAACCCCTTCAAAACCCCAAGAAAGACAGAAATCA	720						
QY	742	AAAGGACTTGGCTCTTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTTTTCAAGAC	801						
DB	721	AAAGGACTTGGCTCTTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTTTTCAAGAC	780						
QY	802	GAGATACCTTAGTCTCAGCAGAGCTGGATGGGATTAATCAAGATAGGATTTACGT	861						
DB	781	GAGATACCTTAGTCTCAGCAGAGCTGGATGGGATTAATCAAGATAGGATTTACGT	840						
QY	862	AAGAAATTAATCTTATCGACAAGAACCCATAGCATCCAACTCTTCTGTACCCAGGT	921						
DB	841	AAGAAATTAATCTTATCGACAAGAACCCATAGCATCCAACTCTTCTGTACCCAGGT	900						
QY	922	AGCAGCACTCGAAACCTTGGATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA	981						
DB	901	AGCAGCACTCGAAACCTTGGATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA	960						
QY	982	TTTGTAAATGACAGACGATTAACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT	1041						
DB	961	TTTGTAAATGACAGACGATTAACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT	1020						
QY	1042	CCAGTGGCTATTTTCAATGGACACCCAGAACTCTACCTTTTATGTAATCCAGCCTTAGT	1101						
QY	1021	CCAGTGGCTATTTTCAATGGACACCCAGAACTCTACCTTTTATGTAATCCAGCCTTAGT	1080						
QY	1102	CCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAATGAAGCTGCTACATATGAAGGTC	1161						
DB	1081	CCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAATGAAGCTGCTACATATGAAGGTC	1140						
QY	1162	TCCACACCCCTGGCAACCTCTACTGTGCTCTCTGGGTCAATTTCTCAAGAGGTCACTCTCTG	1221						
DB	1141	TCCACACCCCTGGCAACCTCTACTGTGCTCTCTGGGTCAATTTCTCAAGAGGTCACTCTCTG	1200						
QY	1222	TGCTGGTGTCCATCTGACTTCAAAAAGATGCTACTGTTCTGTGATGAACAATACATAAAA	1281						
DB	1201	TGCTGGTGTCCATCTGACTTCAAAAAGATGCTACTGTTCTGTGATGAACAATACATAAAA	1260						
QY	1282	ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAAACAGGAGGTGATAAACTTTTCCACGGTG	1341						
DB	1261	ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAAACAGGAGGTGATAAACTTTTCCACGGTG	1320						
QY	1342	GGTTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGCCCTAGTAAACAGTAGTAGC	1401						
DB	1321	GGTTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGCCCTAGTAAACAGTAGTAGC	1380						
QY	1402	CAGAGTACTCTCTGCAAAAGCCCCCAGGGTAAAGTGCAATCCAATTTTCCCGCTCA	1461						
DB	1381	CAGAGTACTCTCTGCAAAAGCCCCCAGGGTAAAGTGCAATCCAATTTTCCCGCTCA	1440						
QY	1462	TCCGAGCTTGTGCCCCCAAGCTGTGAGAGCTCCCTCTCTTCAAAATCTCTCTAGC	1521						
DB	1441	TCCGAGCTTGTGCCCCCAAGCTGTGAGAGCTCCCTCTCTTCAAAATCTCTCTAGC	1500						
QY	1522	TTCTCTATTAATAACCTCTCTGCAAGGCCCGGTCTCCCATCAACAGAAAGGCTCTGTCT	1581						
DB	1501	TTCTCTATTAATAACCTCTCTGCAAGGCCCGGTCTCCCATCAACAGAAAGGCTCTGTCT	1560						
QY	1582	TCCTCCGTCTCTCCCAAGCCACCTTCACTTTTCAAGATGTGATTAAGAACTGGGTGACC	1641						
DB	1561	TCCTCCGTCTCTCCCAAGCCACCTTCACTTTTCAAGATGTGATTAAGAACTGGGTGACC	1620						
QY	1642	CGAACACCTTCTCATCACACCCCATCATCTCCCTGCTCGGAGACCAAGATCATGTCT	1701						
DB	1621	CGAACACCTTCTCATCACACCCCATCATCTCCCTGCTCGGAGACCAAGATCATGTCT	1680						
QY	1702	CCGAGAAAAGCCCTTATTTCTGTGAGCCAGAGGTCACTCCCAAGCAGAGGCTTGCTCTGAG	1761						
DB	1681	CCGAGAAAAGCCCTTATTTCTGTGAGCCAGAGGTCACTCCCAAGCAGAGGCTTGCTCTGAG	1740						
QY	1762	TCTAGAAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAAG	1821						
DB	1741	TCTAGAAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAAG	1800						
QY	1822	TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTGATGGCAAGTTGAAATCTTCAATTTG	1881						
DB	1801	TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTGATGGCAAGTTGAAATCTTCAATTTG	1860						
QY	1882	GATCTGCTGCTGCTTGTGTTAAACAGGAAGACCTTTAGTAAGGACTCTCTAGGTCTCTACC	1941						
DB	1861	GATCTGCTGCTGCTTGTGTTAAACAGGAAGACCTTTAGTAAGGACTCTCTAGGTCTCTACC	1920						
QY	1942	AAATCAAGCAAAAATGAAGGAGCTGGTACCAAGTATCTCAGAGCCCTCCGTCTCTATCAGT	2001						
DB	1921	AAATCAAGCAAAAATGAAGGAGCTGGTACCAAGTATCTCAGAGCCCTCCGTCTCTATCAGT	1980						
QY	2002	CCGTATGCTTCAGAAAGCTGTGGAAACGCTACCTCTTCTTGTGAGACCTTGTGGAGAGGG	2061						
DB	1981	CCGTATGCTTCAGAAAGCTGTGGAAACGCTACCTCTTCTTGTGAGACCTTGTGGAGAGGG	2040						
QY	2062	TCTGAAATGGTAGGCAAGAGAAATAGTTTCCCAAGAGATAAAACTGTGTTGTGGCCATG	2121						
DB	2041	TCTGAAATGGTAGGCAAGAGAAATAGTTTCCCAAGAGATAAAACTGTGTTGTGGCCATG	2100						
QY	2122	GCAGCAAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTCTATCCAGACACCCCAAT	2181						

Db	2101	GCAGCCAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCGGTCATCCAGACACACCCCAAT	2160
Qy	2182	TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACCATCAACGCCAGCTCCATG	2241
Db	2161	TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACCATCAACGCCAGCTCCATG	2220
Qy	2242	AGGAAATCTGCACATACCTCCATAGAAAGTCCAGGAGGACTTCTGTGGTCTCTGAAAC	2301
Db	2221	AGGAAATCTGCACATACCTCCATAGAAAGTCCAGGAGGACTTCTGTGGTCTCTGAAAC	2280
Qy	2302	TCAACAGAAATATAGATCTCTAATCTAGTGAGTACTGAGCTTTTGGTCCACTTAAACAAAG	2361
Db	2281	TCAACAGAAATATAGATCTCTAATCTAGTGAGTACTGAGCTTTTGGTCCACTTAAACAAAG	2340
Qy	2362	CTGAGCTTTTGGTCCACTTAAACCAAGATGAAATACAAGAGTGAATCTATAACTCTGGTC	2421
Db	2341	CTGAGCTTTTGGTCCACTTAAACCAAGATGAAATACAAGAGTGAATCTATAACTCTGGTC	2400
Qy	2422	TTTAAAGAAAGCTGCCCTTTTTCATTTTATAGACAAATCTTTTCAACGCTGAAATGACCTAA	2481
Db	2401	TTTAAAGAAAGCTGCCCTTTTTCATTTTATAGACAAATCTTTTCAACGCTGAAATGACCTAA	2460
Qy	2482	TCGTGTTCTACTACCATATATATATGATGAGCTTCCGAGGATGAATGCTGTGTTTAAATTT	2541
Db	2461	TCGTGTTCTACTACCATATATATATGATGAGCTTCCGAGGATGAATGCTGTGTTTAAATTT	2520
Qy	2542	TCATAAAGTAAATTTGTCTCACTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT	2601
Db	2521	TCATAAAGTAAATTTGTCTCACTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT	2580
Qy	2602	TCATCTTCTATATATATAGATCCAGTTCATGGAGGCAAAAACAAGTTCTCTGTTA	2661
Db	2581	TCATCTTCTATATATATAGATCCAGTTCATGGAGGCAAAAACAAGTTCTCTGTTA	2640
Qy	2662	TCCTGAAACTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCCCTGTCAA	2721
Db	2641	TCCTGAAACTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCCCTGTCAA	2700
Qy	2722	GGCTGACTGAGAAATCTCTGCTGTAAGACCCCTGGTTCTGTTCTGCCCTCCAACATGTATA	2781
Db	2701	GGCTGACTGAGAAATCTCTGCTGTAAGACCCCTGGTTCTGTTCTGCCCTCCAACATGTATA	2760
Qy	2782	ATTTTATTTGAAATACATAATCTTTTTCATG 2814	
Db	2761	ATTTTATTTGAAATACATAATCTTTTTCATG 2793	
RESULT 10			
AX951786			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
, Best Local Similarity			

Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	22	CGATAACGATTTGTGTGTGAGAGCGCAAGTCGATTTTCTGCTGAACCTTTGGAGGCATT	81
Db	1	CGATAACGATTTGTGTGTGAGAGCGCAAGTCGATTTTCTGCTGAACCTTTGGAGGCATT	60
Qy	82	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG	141
Db	61	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG	120
Qy	142	CTCCGCGCAGCCAGCTTTGGCGTCTCGAGAAATGGATGGTCTTTCACAAATACCTCTTCAA	201
Db	121	CTCCGCGCAGCCAGCTTTGGCGTCTCGAGAAATGGATGGTCTTTCACAAATACCTCTTCAA	180
Qy	202	TCCCTTCTGACTGCTTTATCAGTGCAGTGGTAAATGATGAACAACAATTCTTATGGAGAAACA	261
Db	181	TCCCTTCTGACTGCTTTATCAGTGCAGTGGTAAATGATGAACAACAATTCTTATGGAGAAACA	240
Qy	262	GGAGTCCCAGTTCTCTCTTTTGGATGTACCTTCTTCTGCTGCCAATATGGACATGTA	321
Db	241	GGAGTCCCAGTTCTCTCTTTTGGATGTACCTTCTTCTGCTGCCAATATGGACATGTA	300
Qy	322	CTAGCAGTTGGCAATGAAGAAGGCTTTTTCGATTTGATTAACACAGAAATCAAAAGTTTC	381
Db	301	CTAGCAGTTGGCAATGAAGAAGGCTTTTTCGATTTGATTAACACAGAAATCAAAAGTTTC	360
Qy	382	AGAAAGATGCTTTCAAAGAAATGGATGGCTCACCTGGAAATGCCGCTTTTGACCTGGCTGG	441
Db	361	AGAAAGATGCTTTCAAAGAAATGGATGGCTCACCTGGAAATGCCGCTTTTGACCTGGCTGG	420
Qy	442	GTTCTCGTGAACTTAAACCTTTTACAGCAGCAGTGATCAAAAGCCAAATTTTGGGAC	501
Db	421	GTTCTCGTGAACTTAAACCTTTTACAGCAGCAGTGATCAAAAGCCAAATTTTGGGAC	480
Qy	502	GTAAGAGCTGTGAGCTGATTTGGAACATGCAAAAGGTCAATCAATGCAGCCTCAAGTCAGTT	561
Db	481	GTAAGAGCTGTGAGCTGATTTGGAACATGCAAAAGGTCAATCAATGCAGCCTCAAGTCAGTT	540
Qy	562	GCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGGCAACATTATG	621
Db	541	GCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGGCAACATTATG	600
Qy	622	GTCTGGGATACAGGTGCACAAAGATGGGTTTATATAGGCAAGTGAATCAATCAATCAGT	681
Db	601	GTCTGGGATACAGGTGCACAAAGATGGGTTTATATAGGCAAGTGAATCAATCAATCAGT	660
Qy	682	GGAGTCAATATACCTCAGACAAACCCCTTTCAAACCCCAAGAGAAACAGAAATTC	741
Db	661	GGAGTCAATATACCTCAGACAAACCCCTTTCAAACCCCAAGAGAAACAGAAATTC	720
Qy	742	AAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAAAGTTTACTGTGGTCTCTTTCAAGAC	801
Db	721	AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTTTACTGTGGTCTCTTTCAAGAC	780
Qy	802	GAGAAATACCTTAGTCTCAGCAGGAGCTGTGATGGGATTAATCAAGTATGGATTTACGT	861
Db	781	GAGAAATACCTTAGTCTCAGCAGGAGCTGTGATGGGATTAATCAAGTATGGATTTACGT	840
Qy	862	AAGAAATATACCTGTATTCGACAAAGAACCCCATAGCATCCAAAGTCTTTCTGTACCCAGGT	921
Db	841	AAGAAATATACCTGTATTCGACAAAGAACCCCATAGCATCCAAAGTCTTTCTGTACCCAGGT	900
Qy	922	AGCAGACTCGAAAACTTTGGGATTTCAAGTCTGATTTTGGATTTCACTGGCTCTACTTTTA	981
Db	901	AGCAGACTCGAAAACTTTGGGATTTCAAGTCTGATTTTGGATTTCACTGGCTCTACTTTTA	960
Qy	982	TTTGTCTAATTTGCACAGCGATAACATCTACATGTTTAAATATGACTGGGTTCAAGCTTCT	1041
Db	961	TTTGTCTAATTTGCACAGCGATAACATCTACATGTTTAAATATGACTGGGTTCAAGCTTCT	1020
Qy	1042	CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT	1101
Db	1021	CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT	1080


```
QY 1102 CCAGATGACCAAGTTTAACTGCTCAAGTGAAGCTGCTTACATATGGAAGGTC 1161
Db 1081 CCAGATGACCAAGTTTAACTGCTCAAGTGAAGCTGCTTACATATGGAAGGTC 1140
QY 1162 TCCAACCTGCGCAACCTCTCACTGCTGCTCTGGGTCAATCTCAAGAGGTCAAGCTCTGTG 1221
Db 1141 TCCAACCTGCGCAACCTCTCACTGCTGCTCTGGGTCAATCTCAAGAGGTCAAGCTCTGTG 1200
QY 1222 TGCTGGTCCCATCTGACTTCAAAAGATTTGCTTACCTGCTCTGATGACAAATACACTAAA 1281
Db 1201 TGCTGGTCCCATCTGACTTCAAAAGATTTGCTTACCTGCTCTGATGACAAATACACTAAA 1260
QY 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAACTTTTCCACGGTG 1341
Db 1261 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAACTTTTCCACGGTG 1320
QY 1342 GGTGGGCTCTCAGAGAAAGAGTCAAGACTCTGGCTAGTAAACAGTAACAGTAGC 1401
Db 1321 GGTGGGCTCTCAGAGAAAGAGTCAAGACTCTGGCTAGTAAACAGTAACAGTAGC 1380
QY 1402 CAGAGTACTCTGCAAAAGCCCCCAGGTAAGTGCATCCATCCAAATCTTCCCGGTCA 1461
Db 1381 CAGAGTACTCTGCAAAAGCCCCCAGGTAAGTGCATCCATCCAAATCTTCCCGGTCA 1440
QY 1462 TCCGAGCTTTGGCCCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATACTCTTACG 1521
Db 1441 TCCGAGCTTTGGCCCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATACTCTTACG 1500
QY 1522 TTCTCTATTAAACCTCTCTGCAAGGCCGGTCTCCCATCAACAGAGAGGCTCTGTG 1581
Db 1501 TTCTCTATTAAACCTCTCTGCAAGGCCGGTCTCCCATCAACAGAGAGGCTCTGTG 1560
QY 1582 TCCTCGGTCTTCCCAAGCCACTTCACTTTTCAAGATGTGATTTAGAACTGGGTGACC 1641
Db 1561 TCCTCGGTCTTCCCAAGCCACTTCACTTTTCAAGATGTGATTTAGAACTGGGTGACC 1620
QY 1642 CGAACACCTTCTCATCACCCATCACTCCACCTGCTTGGAGACCAAGATCATGTCT 1701
Db 1621 CGAACACCTTCTCATCACCCATCACTCCACCTGCTTGGAGACCAAGATCATGTCT 1680
QY 1702 CCGAGAAAGCCCTTATTCCTGTGTAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1761
Db 1681 CCGAGAAAGCCCTTATTCCTGTGTAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1740
QY 1762 TCTAGAAATAGATGTAAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAG 1821
Db 1741 TCTAGAAATAGATGTAAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAG 1800
QY 1822 TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTGTGATGCGCAAGTTGAAATCTTCATTTG 1881
Db 1801 TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTGTGATGCGCAAGTTGAAATCTTCATTTG 1860
QY 1882 GATCTGTGCTGCTTGTGTTAAACAGAGAACCTTTAGTAAGACTCTCTAGGTCCTTACC 1941
Db 1861 GATCTGTGCTGCTTGTGTTAAACAGAGAACCTTTAGTAAGACTCTCTAGGTCCTTACC 1920
QY 1942 AAATCAAGCAAAATTTGAAGAGCTGGTACCATATCTCAGAGCCTCGGTCTCTTATCAGT 2001
Db 1921 AAATCAAGCAAAATTTGAAGAGCTGGTACCATATCTCAGAGCCTCGGTCTCTTATCAGT 1980
QY 2002 CCGTATGCTTCAGAAAGCTGTGAACGCTTACTCTTCTTGTGAGACCTTGTGGAGAGG 2061
Db 1981 CCGTATGCTTCAGAAAGCTGTGAACGCTTACTCTTCTTGTGAGACCTTGTGGAGAGG 2040
QY 2062 TCTGAAATGGTAGGCAAGAGATAGTTTCCCGAGAGATAAAATCGGTGTGTGGCCATG 2121
Db 2041 TCTGAAATGGTAGGCAAGAGATAGTTTCCCGAGAGATAAAATCGGTGTGTGGCCATG 2100
QY 2122 GCAGCAAAACGGAAGGCTGAGAAATCCATCTCCAGAACTCCGTCTATCCAGACACCCAAAT 2181
Db 2101 GCAGCAAAACGGAAGGCTGAGAAATCCATCTCCAGAACTCCGTCTATCCAGACACCCAAAT 2160
```

```
QY 2182 TCCAGGAGACAGAGCGGAAAGACATTCGAAGCCCGGTCAACATCAGCCAGCTCCATG 2241
Db 2161 TCCAGGAGACAGAGCGGAAAGACATTCGAAGCCCGGTCAACATCAGCCAGCTCCATG 2220
QY 2242 AGGAAATCTGCACATATCTCCATAGAAAGTCCAGAGGAGCTTCTGTGGTCCCTGAAAC 2301
Db 2221 AGGAAATCTGCACATATCTCCATAGAAAGTCCAGAGGAGCTTCTGTGGTCCCTGAAAC 2280
QY 2302 TCAACAGAAATATAGATTTCTAATCTGAGTGAATTTACTGAGCTTTGGTCCACTAAAAACAAG 2361
Db 2281 TCAACAGAAATATAGATTTCTAATCTGAGTGAATTTACTGAGCTTTGGTCCACTAAAAACAAG 2340
QY 2362 CTGAGCTTTGGTCCACTAAAAACAAGATGAAAAATAAAGAGTACTCTATAACTCTGTG 2421
Db 2341 CTGAGCTTTGGTCCACTAAAAACAAGATGAAAAATAAAGAGTACTCTATAACTCTGTG 2400
QY 2422 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2481
Db 2401 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2460
QY 2482 TCTGTTCTACTACCATATGTATATGAGCTTCCCGAGGATGAATGCTGTGTTTAAAT 2541
Db 2461 TCTGTTCTACTACCATATGTATATGAGCTTCCCGAGGATGAATGCTGTGTTTAAAT 2520
QY 2542 TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATPAT 2601
Db 2521 TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATPAT 2580
QY 2602 TCATCTTCTCTAATTAATGATCCAGCTTCAGTGGAGGCAAAAAACAAGTTTCTTGTATA 2661
Db 2581 TCATCTTCTCTAATTAATGATCCAGCTTCAGTGGAGGCAAAAAACAAGTTTCTTGTATA 2640
QY 2662 TCCGTAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGCCCTGTGAA 2721
Db 2641 TCCGTAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGCCCTGTGAA 2700
QY 2722 GGCTGACTGAGAAATCCCTCTGCTGAAGACCCCTGTTCTGTTCGCCCTCCAAACATGATA 2781
Db 2701 GGCTGACTGAGAAATCCCTCTGCTGAAGACCCCTGTTCTGTTCCTCCAAACATGATA 2760
QY 2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793
```

```
RESULT 11
AX960064 LOCUS 4221 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 85 from Patent EP1349104.
ACCESSION AX960064
VERSION AX960064.1 GI:40880290
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Jatkoe, T.
TITLE Method of selecting a portfolio of markers for use in a diagnostic
application
JOURNAL Patent: EP 1349104-A 85 01-OCT-2003;
Ortho-Clinical Diagnostics, Inc. (US)
FEATURES
source
1. 4221
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

ORIGIN

Query Match 98.5%; Score 2789.8; DB 2; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 CGATAACGATTTGTTGTTGAGAGCGCAACGTGCGATTTCTGCTGAACCTTGGAGGCATT 81
Db 1 CGATAACGATTTGTTGTTGAGAGCGCAACGTGCGATTTCTGCTGAACCTTGGAGGCATT 60
QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGAGCCCTGATGCTCTTCAATTCGGTG 141
Db 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGAGCCCTGATGCTCTTCAATTCGGTG 120
QY 142 CTCGCCAGCCCAAGCTTGGCGTCTGAGAAATGGATGGTCTTCAATACCCCTTTCAA 201
Db 121 CTCGCCAGCCCAAGCTTGGCGTCTGAGAAATGGATGGTCTTCAATACCCCTTTCAA 180
QY 202 TCCCTTCTGACTGCTTATCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGAAACA 261
Db 181 TCCCTTCTGACTGCTTATCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGAAACA 240
QY 262 GGAGTCCCAAGTTCCTCTTTTGGATGATACCTTCTCTCTGCTCCCAATATGGAACATGTA 321
Db 241 GGAGTCCCAAGTTCCTCTTTTGGATGATACCTTCTCTCTGCTCCCAATATGGAACATGTA 300
QY 322 CTAGCAGTTGCCAATGGAAGAGCTTTGTTGATTTGATATAACACAGAATCAAAAGTTTC 381
Db 301 CTAGCAGTTGCCAATGGAAGAGCTTTGTTGATTTGATATAACACAGAATCAAAAGTTTC 360
QY 382 AGAAAGAGTGTCTCAAGAAATGGATGGCTCACTGGAATGGCTTTTGAACCTGGCTGG 441
Db 361 AGAAAGAGTGTCTCAAGAAATGGATGGCTCACTGGAATGGCTTTTGAACCTGGCTGG 420
QY 442 GTTCTGTTGAACTTAACTTTGATACAGCAGCAGGTGATCAACACGCCAAATTTTGGGAC 501
Db 421 GTTCTGTTGAACTTAACTTTGATACAGCAGCAGGTGATCAACACGCCAAATTTTGGGAC 480
QY 502 GTAAAAGCTGTGAGCTGATTTGGAACATGCCAAAGTTCATCAATGAGGCTCAAGTCAGTT 561
Db 481 GTAAAAGCTGTGAGCTGATTTGGAACATGCCAAAGTTCATCAATGAGGCTCAAGTCAGTT 540
QY 562 GCCTTTTCTAAGTTTGAAGAACTGATTTCTGTACGGGTGGAGATGGCAACATTATG 621
Db 541 GCCTTTTCTAAGTTTGAAGAACTGATTTCTGTACGGGTGGAGATGGCAACATTATG 600
QY 622 GTCTGGGATACAGGTGCAACAAAAGATGGTTTATAGGCAAGTGAATCAAAATCAGT 681
Db 601 GTCTGGGATACAGGTGCAACAAAAGATGGTTTATAGGCAAGTGAATCAAAATCAGT 660
QY 682 GGAGCTCACAATPACTCAGACAAGCAAAACCCCTTCAAAAACCCAGAGAAGAAATCA 741
Db 661 GGAGCTCACAATPACTCAGACAAGCAAAACCCCTTCAAAAACCCAGAGAAGAAATCA 720
QY 742 AAAGGACTTGCTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTCAGAC 801
Db 721 AAAGGACTTGCTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTCAGAC 780
QY 802 GAGAAATACCTAGTCTCAGCAGAGCTGTGGATGGGTAATCAAAAGTATGGGATTTAGT 861
Db 781 GAGAAATACCTAGTCTCAGCAGAGCTGTGGATGGGTAATCAAAAGTATGGGATTTAGT 840
QY 862 AAGAAATTAATPACTGTATTCGACAAGAACCCATPAGCATCCAAGTCTTTCTGTACCAGGT 921
Db 841 AAGAAATTAATPACTGTATTCGACAAGAACCCATPAGCATCCAAGTCTTTCTGTACCAGGT 900
QY 922 AGCAGACTCGAAATCTGGATTAATCAAGTCTGATTTTGGATTCACCTGGCTCTACTTTA 981
Db 901 AGCAGACTCGAAATCTGGATTAATCAAGTCTGATTTTGGATTCACCTGGCTCTACTTTA 960
QY 982 TTTGCTTAATTCAGACAGCAATCAATCTACATGTTTAAATATGACCTGGGTTGAAGACTTCT 1041
Db 961 TTTGCTTAATTCAGACAGCAATCAATCTACATGTTTAAATATGACCTGGGTTGAAGACTTCT 1020
QY 1042 CCAGTGGCTATTTTCAATGGACACCAAGAACTTACCTTTTATGTAAAAATCCAGCCTTAGT 1101
Db 1021 CCAGTGGCTATTTTCAATGGACACCAAGAACTTACCTTTTATGTAAAAATCCAGCCTTAGT 1080

QY 1102 CCAGATGACCAAGTTTTTAGTCAGTGGCTCAAGTGAAGCTGCTACATATGAAGGTC 1161
Db 1081 CCAGATGACCAAGTTTTTAGTCAGTGGCTCAAGTGAAGCTGCTACATATGAAGGTC 1140
QY 1162 TCCACACCCCTGGCAACCTCTCTACTTGTCTCTCTGGGTCAATTTCTCAAGAGGTCACTCTGTG 1221
Db 1141 TCCACACCCCTGGCAACCTCTCTACTTGTCTCTCTGGGTCAATTTCTCAAGAGGTCACTCTGTG 1200
QY 1222 TGCTGGTGTCCATCTGACTTTCACAAAGATTTGCTACTCTTGTCTGATGACAAATACACTAAAA 1281
Db 1201 TGCTGGTGTCCATCTGACTTTCACAAAGATTTGCTACTCTTGTCTGATGACAAATACACTAAAA 1260
QY 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACACAGAGGTGATAAATTTTCAACGGTG 1341
Db 1261 ATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACACAGAGGTGATAAATTTTCAACGGTG 1320
QY 1342 GGTTTGGSCCTCTCAGNAGAAAAAGAGTCAAGACCTGGCCCTAGTAAACAGTAAACGAGTAGC 1401
Db 1321 GGTTTGGSCCTCTCAGNAGAAAAAGAGTCAAGACCTGGCCCTAGTAAACAGTAAACGAGTAGC 1380
QY 1402 CAGAGTACTCTCTGCAAAAGCCCCCAGGGTAAAGTGCAATCCATCCAATTTCTTCCCGCTCA 1461
Db 1381 CAGAGTACTCTCTGCAAAAGCCCCCAGGGTAAAGTGCAATCCATCCAATTTCTTCCCGCTCA 1440
QY 1462 TCCGACAGTTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCAAATACTCTCTACG 1521
Db 1441 TCCGACAGTTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCAAATACTCTCTACG 1500
QY 1522 TTTCTTATTAACCTCTCTGTCNAGGCCGGTCTCCCATCAACAGAGAAGGCTCTGTCT 1581
Db 1501 TTTCTTATTAACCTCTCTGTCNAGGCCGGTCTCCCATCAACAGAGAAGGCTCTGTCT 1560
QY 1582 TCTCTCGCTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTGATAGAACTGGGTGACC 1641
Db 1561 TCTCTCGCTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTGATAGAACTGGGTGACC 1620
QY 1642 CGAACACCTTCTCATCACCAACCCNCTCACTCCACTGCTCGGAGACCAAGATCATGTCT 1701
Db 1621 CGAACACCTTCTCATCACCAACCCNCTCACTCCACTGCTCGGAGACCAAGATCATGTCT 1680
QY 1702 CCGAGAAAAAGCCCTTATTTCTGTGAGCCAGAAGTCATCCCAAGCAGAGGCTTGTCTGAG 1761
Db 1681 CCGAGAAAAAGCCCTTATTTCTGTGAGCCAGAAGTCATCCCAAGCAGAGGCTTGTCTGAG 1740
QY 1762 TCTAGAAATAGAGTAAAGAGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAG 1821
Db 1741 TCTAGAAATAGAGTAAAGAGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAG 1800
QY 1822 TGTGTGAAGAGTGTCTAACTGTGACTGAGCTTGATGGCCAAAGTTGAAAAATCTTCAATTG 1881
Db 1801 TGTGTGAAGAGTGTCTAACTGTGACTGAGCTTGATGGCCAAAGTTGAAAAATCTTCAATTG 1860
QY 1882 GATCTGTCTGCTTGTGTGTAACAGGAAGACCTTAGTAAGGACTCTCTAGGTCCTTACC 1941
Db 1861 GATCTGTCTGCTTGTGTGTAACAGGAAGACCTTAGTAAGGACTCTCTAGGTCCTTACC 1920
QY 1942 AAATCAAGCAAAATTTGAAGAGCTGTGTACAGTATCTCAGAGCCCTCGTCTCTTACGT 2001
Db 1921 AAATCAAGCAAAATTTGAAGAGCTGTGTACAGTATCTCAGAGCCCTCGTCTCTTACGT 1980
QY 2002 CCGTATGCTTCAGAAAGCTGTGGAAACGCTACTCTTCTTTTGGAGACCTTGTGGAGAGGG 2061
Db 1981 CCGTATGCTTCAGAAAGCTGTGGAAACGCTACTCTTCTTTTGGAGACCTTGTGGAGAGGG 2040
QY 2062 TCTGAAATGGTAGGCAAGAGAAATAGTTTCCCAGAGAAATAAAAACTGGTGTGTGGCCATG 2121
Db 2041 TCTGAAATGGTAGGCAAGAGAAATAGTTTCCCAGAGAAATAAAAACTGGTGTGTGGCCATG 2100
QY 2122 GCAGCAAAACGGAGGCTGAGAAATCCATCTCCAGAAAGTCCGTCTATCCACACACCCCAAT 2181
Db 2101 GCAGCAAAACGGAGGCTGAGAAATCCATCTCCAGAAAGTCCGTCTATCCACACACCCCAAT 2160
QY 2182 TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAACGCCCAGCTTCCATG 2241

```
Db 2161 TCAGGAGACAGCGGGAAGACATTTGCCAAGCCGGTCACCATCAGGCCAGCTCCATG 2220
Qy 2242 AGGAAATCTGCACATACCTCCATAGAAAGTCCAGGAGAGCTTCTGTGGTCTGAAACAC 2301
Db 2221 AGGAAATCTGCACATACCTCCATAGAAAGTCCAGGAGAGCTTCTGTGGTCTGAAACAC 2280
Qy 2302 TCAACAGAAATTAGAATCTTAATCTCAGTGAGTTACTGTAGCTTTGGTCCACATAAACACAG 2361
Db 2281 TCACAGAAATTAGAATCTTAATCTCAGTGAGTTACTGTAGCTTTGGTCCACATAAACACAG 2340
Qy 2362 CTGAGCTTTGGTCCATAAACAAGATGAAAAATACAAGAGTGACTCTATAACTCTGGTC 2421
Db 2341 CTGAGCTTTGGTCCATAAACAAGATGAAAAATACAAGAGTGACTCTATAACTCTGGTC 2400
Qy 2422 TTAAAGAAAGCTGCCCTTTTCATTTTGTAGACAAAATCTTTTCAACGCTGAAATGTACCTAA 2481
Db 2401 TTAAAGAAAGCTGCCCTTTTCATTTTGTAGACAAAATCTTTTCAACGCTGAAATGTACCTAA 2460
Qy 2482 TCTGGTTCTACTACCATATATGTATATGACGTTCCGAGGATGAATGCTGTGTTAAAT 2541
Db 2461 TCTGGTTCTACTACCATATATGTATATGACGTTCCGAGGATGAATGCTGTGTTAAAT 2520
Qy 2542 TCATAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATAT 2601
Db 2521 TCATAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATAT 2580
Qy 2602 TCATCTCTCTATAATAATAGCATCCAGTTCATGGAGGCAAAAACAAGTTCTTGTTA 2661
Db 2581 TCATCTCTCTATAATAATAGCATCCAGTTCATGGAGGCAAAAACAAGTTCTTGTTA 2640
Qy 2662 TCCTGAAACTTCTATGCTCAGTGGAAGTATCTGCCAGCCACAGCATGAGCCGTGTGA 2721
Db 2641 TCTGAAACTTCTATGCTCAGTGGAAGTATCTGCCAGCCACAGCATGAGCCGTGTGA 2700
Qy 2722 GGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCTGTCTCCACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCTGTCTCCACATGTATA 2760
Qy 2782 ATTTTATTGAAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTGAAATACATAATCTTTTCACTATG 2793
```

```
RESULT 12
CS033737
LOCUS CS033737 4221 bp DNA linear PAT 10-MAR-2005
DEFINITION Sequence 3243 from Patent WO2005016962.
ACCESSION CS033737
VERSION CS033737.1 GI:60732651
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Abbas, A., Clark, H., Ouyang, W., Williams, M.P., Wood, W.I. and Wu, T.D.
TITLE Compositions and methods for the treatment of immune related
diseases
JOURNAL Patent: WO 2005016962-A 3243 24-FEB-2005;
Genentech, Inc. (US)
FEATURES
source
1. .4221
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

ORIGIN

```
Query Match 98.5%; Score 2789.8; DB 2; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 22 CGATAACGATTTGTTGTTGAGAGCGCAACGTCGATTTCTGTGAACTTTGGAGGCATT 81
Db 1 CGATAACGATTTGTTGTTGAGAGCGCAACGTCGATTTCTGTGAACTTTGGAGGCATT 60
Qy 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 141
Db 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 120
Qy 142 CTCGCCAGCCAGCTTGGGCTCTGAGAAATGATGGTCTTCAATACCTCTTCAA 201
Db 121 CTCGCCAGCCAGCTTGGGCTCTGAGAAATGATGGTCTTCAATACCTCTTCAA 180
Qy 202 TCCCTTCTGACTGGTTATCAGTGCAAGTAAATGATGAACACACTTCTTATGGAGAAACA 261
Db 181 TCCCTTCTGACTGGTTATCAGTGCAAGTAAATGATGAACACACTTCTTATGGAGAAACA 240
Qy 262 GGAGTCCAGTTCTCTCTTTTGGATGATACCTTCTCTGCTCCCAATATGGAACATGTA 321
Db 241 GGAGTCCAGTTCTCTCTTTTGGATGATACCTTCTCTGCTCCCAATATGGAACATGTA 300
Qy 322 CTAGCAGTTGCCAATGAAAGAGGCTTTGTTGATTTGATATACACAGAAATCAAAAGTTTC 381
Db 301 CTAGCAGTTGCCAATGAAAGAGGCTTTGTTGATTTGATATACACAGAAATCAAAAGTTTC 360
Qy 382 AGAAGAAAGTCTTCAAAAGAAATGGATGGCTCACTGGAAATGCCGCTTTTGACCTGGCCTGG 441
Db 361 AGAAGAAAGTCTTCAAAAGAAATGGATGGCTCACTGGAAATGCCGCTTTTGACCTGGCCTGG 420
Qy 442 GTTCTGTGTGAATTTAAACTTTGTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC 501
Db 421 GTTCTGTGTGAATTTAAACTTTGTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC 480
Qy 502 GTAAAGCTGTGAGCTGATTTGGAACATGCAAGGTCATCAATGCAGCTCAAGTCAGTT 561
Db 481 GTAAAGCTGTGAGCTGATTTGGAACATGCAAGGTCATCAATGCAGCTCAAGTCAGTT 540
Qy 562 GCCTTTTCTAAGTTTGAGAAAGCTGTTATCTGTACGGTGGAAGAGATGGCAACATTATG 621
Db 541 GCCTTTTCTAAGTTTGAGAAAGCTGTTATCTGTACGGTGGAAGAGATGGCAACATTATG 600
Qy 622 GTCTGGGATACAGGTGCAACAAAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT 681
Db 601 GTCTGGGATACAGGTGCAACAAAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT 660
Qy 682 GGAGCTCAATATCCTGACAGACGCAACCCCTTCAAAACCCCAAGAAACAGAAATTCATCA 741
Db 661 GGAGCTCAATATCCTGACAGACGCAACCCCTTCAAAACCCCAAGAAACAGAAATTCATCA 720
Qy 742 AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTCTTTCAAGAC 801
Db 721 AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTCTTTCAAGAC 780
Qy 802 GAGAACTCTTAGTCTCAGCAGGAGCTGTGATGGGATTAATCAAAAGTATGGATTTACGT 861
Db 781 GAGAACTCTTAGTCTCAGCAGGAGCTGTGATGGGATTAATCAAAAGTATGGATTTACGT 840
Qy 862 AAGAAATATATGCTTATGACAGAACCCATAGCATCCAGTCTTCTCTGACCCAGGT 921
Db 841 AAGAAATATATGCTTATGACAGAACCCATAGCATCCAGTCTTCTCTGACCCAGGT 900
Qy 922 AGCAGCTCGAAAACTTTGGATATTTCAAGTCTGATTTTGGATTTCCAGCTGGCTCTACTTTTA 981
Db 901 AGCAGCTCGAAAACTTTGGATATTTCAAGTCTGATTTTGGATTTCCAGCTGGCTCTACTTTTA 960
Qy 982 TTTGCTAAATGCACAGACGATAACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT 1041
Db 961 TTTGCTAAATGCACAGACGATAACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT 1020
Qy 1042 CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1101
Db 1021 CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1080
Qy 1102 CCAGTATGACAGTTTATGTCAGTGGCTCAAGTATGAAGCTGCCTACATATGGAAGGTC 1161
```

```
|||||
1081 CCAGATGACCAGTTTCTAGTGGCTCAAGTGATGAAGCTGCCTACATATGGAGGTC 1140
|||||
1162 TCCACACCGTGGCAACCTCTACTGTGCTCTGGGTCAATCTCAAGAGGTCAAGCTCTGTG 1221
|||||
1141 TCCACACCGTGGCAACCTCTACTGTGCTCTGGGTCAATCTCAAGAGGTCAAGCTCTGTG 1200
|||||
1222 TCGTGGTCCATCTGACTTCAAAAGATTGCTACTGTTCTGATGACAATACACTAAAA 1281
|||||
1201 TGCCTGGTCCATCTGACTTCAAAAGATTGCTACTGTTCTGATGACAATACACTAAAA 1260
|||||
1282 ATCTCGCGCTTCAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCACGGTG 1341
|||||
1261 ATCTCGCGCTTCAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCACGGTG 1320
|||||
1342 GGTGGGCGCTCTCAGAGAAAAAGAGTCAAGACCTGGCTAGTAACAGTAACAGTAGC 1401
|||||
1321 GGTGGGCGCTCTCAGAGAAAAAGAGTCAAGACCTGGCTAGTAACAGTAACAGTAGC 1380
|||||
1402 CAGAGTACTCTCGCCAAAGCCCCCAGGGTAAAGTGGAATCCATCCAAATCTTCCCGGTCA 1461
|||||
1381 CAGAGTACTCTCGCCAAAGCCCCCAGGGTAAAGTGGAATCCATCCAAATCTTCCCGGTCA 1440
|||||
1462 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCTACG 1521
|||||
1441 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCTACG 1500
|||||
1522 TTCTCTATTAAACCTCTCTGCCAAGCCCGGTCTCCCATCAACAGAGAGGCTCTGTC 1581
|||||
1501 TTCTCTATTAAACCTCTCTGCCAAGCCCGGTCTCCCATCAACAGAGAGGCTCTGTC 1560
|||||
1582 TCCTCGGTCTCTCCAAAGCACCTTCATCTTTCAAGATGTGATAGAACTGGGTGACC 1641
|||||
1561 TCCTCGGTCTCTCCAAAGCACCTTCATCTTTCAAGATGTGATAGAACTGGGTGACC 1620
|||||
1642 CGAAACACCTTCTCATCACCCATCACTCCACCTGCTTGGAGACCAAGATCATGTCT 1701
|||||
1621 CGAAACACCTTCTCATCACCCATCACTCCACCTGCTTGGAGACCAAGATCATGTCT 1680
|||||
1702 CCGAGAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGCTCTGAG 1761
|||||
1681 CCGAGAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGCTCTGAG 1740
|||||
1762 TCTAGAAATAGAGTAAAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1821
|||||
1741 TCTAGAAATAGAGTAAAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1800
|||||
1822 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGGCCAAAGTTGAAAACTCTCATTTG 1881
|||||
1801 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGGCCAAAGTTGAAAACTCTCATTTG 1860
|||||
1882 GATCTGTGCTTGCTTGCTGTTAACCAGGAAGACCTTAGTAAGGACTCTCTAGGTCCTACC 1941
|||||
1861 GATCTGTGCTTGCTTGCTGTTAACCAGGAAGACCTTAGTAAGGACTCTCTAGGTCCTACC 1920
|||||
1942 AAATCAAGCAAAATTTGAAGAGCTGGTACCAGTATCTCAGAGCCTCCGTCTCCTATCAGT 2001
|||||
1921 AAATCAAGCAAAATTTGAAGAGCTGGTACCAGTATCTCAGAGCCTCCGTCTCCTATCAGT 1980
|||||
2002 CCGTATGCTTCAGAAAGCTGTGAAACGCTTACCTCTTCTTTCAGACCTTGTGGAGAGGG 2061
|||||
1981 CCGTATGCTTCAGAAAGCTGTGAAACGCTTACCTCTTCTTTCAGACCTTGTGGAGAGGG 2040
|||||
2062 TCTGAAATGGTAGGCAAGAGAAATAGTTCCCGAGAGAAATAAAACTGGTGTGTGGCCATG 2121
|||||
2041 TCTGAAATGGTAGGCAAGAGAAATAGTTCCCGAGAGAAATAAAACTGGTGTGTGGCCATG 2100
|||||
2122 GCAGCAAAACGGAAGCTGAGAAATCGATCTCCAGAAAGTCCGTCTATCCAGACACCCAT 2181
|||||
2101 GCAGCAAAACGGAAGCTGAGAAATCGATCTCCAGAAAGTCCGTCTATCCAGACACCCAT 2160
|||||
2182 TCAGGAGACAGCGGGAAGACATTTGCCAAGCCCGGTCAACCATCAAGCCCGAGCTCCCATG 2241
|||||
```

```
Db 2161 TCCAGGAGACAGAGCGGAAAGACATTCGCAAGCCCGGTCCACCATCAAGCCCGAGCTCCATG 2220
Qy 2242 AGGAAATCTGCACATATCTTCCATAGAAAAGTCCAGGAGNCTTCTGTGGTCCCTGAACAC 2301
Db 2221 AGGAAATCTGCACATATCTTCCATAGAAAAGTCCAGGAGGACTTCTGTGGTCCCTGAACAC 2280
Qy 2302 TCACAGAAATTTATAGATTCTTAATCTGAGTGTAGTTACTGAGCTTTGGTCCACTTAAACCAAG 2361
Db 2281 TCAACAGAAATTTATAGATTCTTAATCTGAGTGTAGTTACTGAGCTTTGGTCCACTTAAACCAAG 2340
Qy 2362 CTGAGCTTTGGTCCACTTAAACCAAGATGAAAAATACAAAGTGTGACTCTATAACTCTGGTC 2421
Db 2341 CTGAGCTTTGGTCCACTTAAACCAAGATGAAAAATACAAAGTGTGACTCTATAACTCTGGTC 2400
Qy 2422 TTTAAGAAAGCTGCTCTTTTCAATTTTATAGACAAATATCTTTTCAACGCTGAAATGTACCTAA 2481
Db 2401 TTTAAGAAAGCTGCTCTTTTCAATTTTATAGACAAATATCTTTTCAACGCTGAAATGTACCTAA 2460
Qy 2482 TCTGGTCTTACTACCATTAATGTATATGCAGCTTCCCGAGGATGAATGCTGTGTTTAAATT 2541
Db 2461 TCTGGTCTTACTACCATTAATGTATATGCAGCTTCCCGAGGATGAATGCTGTGTTTAAATT 2520
Qy 2542 TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2601
Db 2521 TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2580
Qy 2602 TCATCTTCTCTATAATTAATGACATCCAGATTCAGAGGCAAAAAACAGTTTCTGTGTA 2661
Db 2581 TCATCTTCTCTATAATTAATGACATCCAGATTCAGAGGCAAAAAACAGTTTCTGTGTA 2640
Qy 2662 TCCGAAACTTCTCATGCTCAGTGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
Db 2641 TCCTGAAACTTCTCATGCTCAGTGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
Qy 2722 GGCTGACTGAGAAATCTCTGCTGAAGACCCCTGGTCTGTGTTCTGCTCCCAACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCTCTGCTGAAGACCCCTGGTCTGTGTTCTGCTCCCAACATGTATA 2760
Qy 2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793
```

RESULT 13
CS042689 LOCUS CS042689 4221 bp DNA linear PAT 22-MAR-2005
DEFINITION Sequence 3243 from Patent WO2005019258.
ACCESSION CS042689
VERSION CS042689.1 GI:61849834
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Abbas, A., Clark, H., Ouyang, W., Williams, P.M., Wood, W.I. and Wu, T.D.
TITLE Compositions and methods for the treatment of immune related diseases
JOURNAL Patent: WO 2005019258-A 3243 03-MAR-2005;
Genentech, Inc. (US)

FEATURES
source Location/Qualifiers
1..4221
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 98.5%; Score 2789.8; DB 2; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 22 CGATAACGATTGTGTGTTGTGAGAGCGCAACGTGCGATTTCTGCTGAACCTTGGAGGCATT 81

1 CGATAACGATTTGTGTGTGAGAGCGCAAGCTGCGATTTCTGCTGAACTTGGAGGCAAT 60
82 TCTACGACATTTCTCTCAGCTGAGGCTTTTCTCCGACCTGATGCTCTTCAATTCGGTG 141
61 TCTACGACATTTCTCTCAGCTGAGGCTTTTCTCCGACCTGATGCTCTTCAATTCGGTG 120
142 CTCGCGCAGCCGAGCTTGGCGTCTCGAGAAATGGATGCTTCAAAATACCTCTTTCAA 201
121 CTCGCGCAGCCGAGCTTGGCGTCTCGAGAAATGGATGCTTCAAAATACCTCTTTCAA 180
202 TCCCTTCTGACTGGTTATCAGTGCAGTGTGTAATGATGAACACATTTCTTATGGAGAAACA 261
181 TCCCTTCTGACTGGTTATCAGTGCAGTGTGTAATGATGAACACACTTCTTATGGAGAAACA 240
262 GGNATCCGAGTTCCTCTTTGGATGTAACCTTCTCTTCTGCTCCCAATATGGAACATGTA 321
241 GGNATCCGAGTTCCTCTTTGGATGTAACCTTCTCTTCTGCTCCCAATATGGAACATGTA 300
322 CTAGCAGTTGCCAATGAAGGCTTTGTTGCAATGTAATGAACACAGAAATCAACAAGTTTC 381
301 CTAGCAGTTGCCAATGAAGGCTTTGTTGCAATGTAATGAACACAGAAATCAACAAGTTTC 360
382 AGAAAGAGTGTCTTCAAAGATGGAATGCGTCACTGGAATGCCGTCTTTGACCTGGCCTGG 441
361 AGAAAGAGTGTCTTCAAAGATGGAATGCGTCACTGGAATGCCGTCTTTGACCTGGCCTGG 420
442 GTTCTCGTGAATTAACCTTTAGCAGAGCAGTGTATCAACAGCCCAAAATTTGGGAC 501
421 GTTCTCGTGAATTAACCTTTAGCAGAGCAGTGTATCAACAGCCCAAAATTTGGGAC 480
502 GTAAAAGCTGTGAGCTGATGGAACATGCAAGGTCATCAATGCAAGCTCAAGTCAGTT 561
481 GTAAAAGCTGTGAGCTGATGGAACATGCAAGGTCATCAATGCAAGCTCAAGTCAGTT 540
562 GCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGTGGAAGATGGCAACATTTATG 621
541 GCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGTGGAAGATGGCAACATTTATG 600
622 GTCTGGGATACAGGTGCAACAAAAGATGGTTTTATAGGCAAGTGAATCAATCAGT 681
601 GTCTGGGATACAGGTGCAACAAAAGATGGTTTTATAGGCAAGTGAATCAATCAGT 660
682 GGAGCTCACAATACCTCAGACAGCAACCCCTTCAAAACCCCAAGAGAAACAGAAATCA 741
661 GGAGCTCACAATACCTCAGACAGCAACCCCTTCAAAACCCCAAGAGAAACAGAAATCA 720
742 AAAGGACTTGTCTCTGTGGATTTCCAGCAAAAGTGTACTGTGCTCTCTTTCAAAGAC 801
721 AAAGGACTTGTCTCTGTGGATTTCCAGCAAAAGTGTACTGTGCTCTCTTTCAAAGAC 780
802 GAGAAATACCTTATCTCAGAGAGCTGTGGATGGGATAATCAAGATGGGATTTACGT 861
781 GAGAAATACCTTATCTCAGAGAGCTGTGGATGGGATAATCAAGATGGGATTTACGT 840
862 AAGAATATACCTTATCTCAGAGAGCTGTGGATGGGATAATCAAGATGGGATTTACGT 921
841 AAGAATATACCTTATCTCAGAGAGCTGTGGATGGGATAATCAAGATGGGATTTACGT 900
922 AGCAGCACTCGAAAATCTGGATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
901 AGCAGCACTCGAAAATCTGGATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
982 TTTGCTTAATGCAACAGCATAACACTACTACTGTTTTAATATGACTGGGTGGAAGCTTCT 1041
961 TTTGCTTAATGCAACAGCATAACACTACTACTGTTTTAATATGACTGGGTGGAAGCTTCT 1020
1042 CCAGTGGCTATTTTCAATGGACCAAGAACTTACCTTTTATGTAATTCGAAGCTTACT 1101
1021 CCAGTGGCTATTTTCAATGGACCAAGAACTTACCTTTTATGTAATTCGAAGCTTACT 1080
1102 CCAGATGACAGTTTTTTAGTCAAGTGGCTCAAGTGAATGAAGCTGCCTTACATATGGAAGGTC 1161

1081 CCAGATGACCAAGTTTTTTAGTCAAGTGGCTCAAGTGAAGCTGCCTACATATGGAAGGTC 1140
1162 TCCACACCTCGCAACCTCTCTACTGTGCTCTCTGGGTCAATTTCTCAAGAGGTCAAGTCTGTG 1221
1141 TCCACACCTCGCAACCTCTCTACTGTGCTCTCTGGGTCAATTTCTCAAGAGGTCAAGTCTGTG 1200
1222 TGCTGGTGTCCATCTGACTTTCACAAAGATGCTACTCTGTTCTTGATGACAATACACTAAAA 1281
1201 TGCTGGTGTCCATCTGACTTTCACAAAGATGCTACTCTGTTCTTGATGACAATACACTAAAA 1260
1282 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAAACAGAGGTGATAAACTTTTCCACGGTG 1341
1261 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAAACAGAGGTGATAAACTTTTCCACGGTG 1320
1342 GGTGGGCGCTCTCAGAAAGAAAAGAGTCAAGACCTGCGCTAGTAACAGTAACAGTAGTAC 1401
1321 GGTGGGCGCTCTCAGAAAGAAAAGAGTCAAGACCTGCGCTAGTAACAGTAACAGTAGTAC 1380
1402 CAGAGTACTCTCTGCCAAAGCCCCCAGAGGTAAAGTGAATCCATCCAAATTTCTTCCCGTCA 1461
1381 CAGAGTACTCTCTGCCAAAGCCCCCAGAGGTAAAGTGAATCCATCCAAATTTCTTCCCGTCA 1440
1462 TCCGAGCTGTGCGCCCAAGCTGTGCTGGAGACCTCTCTCTTCTTCAAAATACCTCTAG 1521
1441 TCCGAGCTGTGCGCCCAAGCTGTGCTGGAGACCTCTCTCTTCTTCAAAATACCTCTAG 1500
1522 TTCTCTATTAATAAATCTCTCTGCCAAGCGCGGTCTCCCATCAACAGAGAGGCTCTGTG 1581
1501 TTCTCTATTAATAAATCTCTCTGCCAAGCGCGGTCTCCCATCAACAGAGAGGCTCTGTG 1560
1582 TCCTCCGTCTCTCCCAAGCCACCTTTCATCTTTCAAGATGTGATTTAGAAAATGGGTGACC 1641
1561 TCCTCCGTCTCTCCCAAGCCACCTTTCATCTTTCAAGATGTGATTTAGAAAATGGGTGACC 1620
1642 CGAACACTTCTCTCATCAACCAATCACTCCACTGCTGTGGAGACCAAGATCATGTCT 1701
1621 CGAACACTTCTCTCATCAACCAATCACTCCACTGCTGTGGAGACCAAGATCATGTCT 1680
1702 CCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTCAG 1761
1681 CCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTCAG 1740
1762 TCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTCTCTGGAGAGTGTGAAAACAAAAG 1821
1741 TCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTCTCTGGAGAGTGTGAAAACAAAAG 1800
1822 TGTGTGAAGAGTGTGTAACCTGTGACTGAGCTTGAATGCCCAAGTTGAAAATCTTCAATTTG 1881
1801 TGTGTGAAGAGTGTGTAACCTGTGACTGAGCTTGAATGCCCAAGTTGAAAATCTTCAATTTG 1860
1882 GATCTGTGCTGCTTGTGCTGTAACCAAGAGACCTTAGTAGGACTCTCTAGGTCCTTACC 1941
1861 GATCTGTGCTGCTTGTGCTGTAACCAAGAGACCTTAGTAGGACTCTCTAGGTCCTTACC 1920
1942 AAATCAAGCAAAAATGGAAGGAGTGTGTAACCTGTGACTGAGCTTGAATGCCCAAGTTGAAAATCTTCAATTTG 2001
1921 AAATCAAGCAAAAATGGAAGGAGTGTGTAACCTGTGACTGAGCTTGAATGCCCAAGTTGAAAATCTTCAATTTG 1980
2002 CCGTATGCTTCCAGAAAGCTGTGGAACGCTACTCTTCTCTTTGAGACCTTGTGGAGAAAGG 2061
1981 CCGTATGCTTCCAGAAAGCTGTGGAACGCTACTCTTCTCTTTGAGACCTTGTGGAGAAAGG 2040
2062 TCTGAAATGGTAGGCAAAAGAGATGTTCCCGAGAGAAATAAAATGTTGTTGGCCATG 2121
2041 TCTGAAATGGTAGGCAAAAGAGATGTTCCCGAGAGAAATAAAATGTTGTTGGCCATG 2100
2122 GCAGCAAAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTCATCTCCAGACACCCCAAT 2181
2101 GCAGCAAAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTCATCTCCAGACACCCCAAT 2160
2182 TCCAGGAGACAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAACGCCCAGCTTCCATG 2241
2161 TCCAGGAGACAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAACGCCCAGCTTCCATG 2220

QY	2242	AGGAAATCTGCACATACCTTCCATAGAAAGTCCAGAGGAGCTTCTGTGGTCTCGAACAC	2301
Db	2221	AGGAAATCTGCACATACCTTCCATAGAAAGTCCAGAGGAGCTTCTGTGGTCTCGAACAC	2280
QY	2302	TCAACAGAAATTATAGATTCTTAATCTGAGTGAAGTACTGAGCTTTGGTCCACCTAAACAAAG	2361
Db	2281	TCAACAGAAATTATAGATTCTTAATCTGAGTGAAGTACTGAGCTTTGGTCCACCTAAACAAAG	2340
QY	2362	CTGAGCTTTGGTCCACTAAACAGATGAAATAACAAGAGTGACTCTATAACTCTGGTC	2421
Db	2341	CTGAGCTTTGGTCCACTAAACAGATGAAATAACAAGAGTGACTCTATAACTCTGGTC	2400
QY	2422	TTTAAGAAAGTCCCTTTTCATTTTATAGACAAATCTTTTCAAGCTGAATGACCTAA	2481
Db	2401	TTTAAGAAAGTCCCTTTTCATTTTATAGACAAATCTTTTCAAGCTGAATGACCTAA	2460
QY	2482	TCGTGTTCTACTACCATATGATGATGAGCTTCCGAGGAGTGAATGCTGTGTTTAAATT	2541
Db	2461	TCGTGTTCTACTACCATATGATGATGAGCTTCCGAGGAGTGAATGCTGTGTTTAAATT	2520
QY	2542	TCATAAGTAAATTTGTCACTAGCATTTTGAATGAATAGTCTTTCACCTTTTAAATAT	2601
Db	2521	TCATAAGTAAATTTGTCACTAGCATTTTGAATGAATAGTCTTTCACCTTTTAAATAT	2580
QY	2602	TCATCTTCTATATAATAGATCCAGTTCATGAGGAGCAAAAACAAGTTCTTGTTA	2661
Db	2581	TCATCTTCTATATAATAGATCCAGTTCATGAGGAGCAAAAACAAGTTCTTGTTA	2640
QY	2662	TCCTGAACTTCTATGCTCAGTGAAGAGTACTGCCAGCCACACATGAGGCCCTGTGAA	2721
Db	2641	TCCTGAACTTCTATGCTCAGTGAAGAGTACTGCCAGCCACACATGAGGCCCTGTGAA	2700
QY	2722	GGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTCTTCTGCTCCCAACATGATA	2781
Db	2701	GGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTCTTCTGCTCCCAACATGATA	2760
QY	2782	ATTTATTGAAATACATAATCTTTTCACTATG	2814
Db	2761	ATTTATTGAAATACATAATCTTTTCACTATG	2793
RESULT 14			
AF195765			
LOCUS			
DEFINITION	Homo sapiens L2DTL protein (L2DTL) mRNA, linear PRI 21-FEB-2000		
ACCESSION	AF195765		
VERSION	AF195765.1 GI:7012713		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1 (bases 1 to 4221)		
AUTHORS	Mueller, R. and Ziegler, B. L.		
TITLE	Identification of L2DTL, a human WD-40 repeat gene homolog of the		
	Drosophila lethal (2) denticleless heat shock gene [1(2)dtl]		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 4221)		
AUTHORS	Mueller, R. and Ziegler, B. L.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-OCT-1999) Kimmel Cancer Institute, Thomas Jefferson		
	University, 233 South 10th Street, Philadelphia, PA 19107-5541, USA		
FEATURES	Location/Qualifiers		
source	1..4221		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
gene	1..4221		
	/gene="L2DTL"		
CDS	103..2295		
	/gene="L2DTL"		

/note="contains WD-40 repeats; similar to the Drosophila lethal (2) denticleless heat shock gene, 1(2)dtl"
/codon_start=1
/product="L2DTL protein"
/protein_id="AAF35182.1"
/db_xref="GI:7012714"
/translation="MLFNSVLROPLGVLNGWSSOYVLOSLLTGYOCSGNDEHTSYG
ETGVVPFGCTFSAPDNMEHVLAVANEFGVRLYNTESSQFRKCKEWAHNAV
DLAWPGELKLVTAAGDTAFWDVAGELIGTCKHQCSLKSVAFSKFAVFTGG
RDGIMVMDTRCNKQDGYRQVQIISGHNATSDQTPSKPKKONSGLSPVDFQS
VTVILFODENTLVSAAGVDGI IKVMDLRKNYTAIRQPIAS KSPLYPGSTRKLGYSS
LILDSGTILFANCTDDNIYMFNMTGLKTSFPAIFNGHONSTFVYKSLSPDDQFLVS
GSDERAATIKVSTPMOPPTVLLGHSEVTSVWCPSDFTKIATCSDNDLTKIWLNR
GLBEKPGDKLSTVGHASOKKESRGLVTVTSOSTPAKAPRVCKNPSNPSAAC
APSCAGLPULPNTPTFSIKTISPAKARSPINRRGSVSSVSPKPSFMSIRNWTIT
PSSSPITTPASETKINSPRKALIPVSOQAACSESRNRVRRVRLDSSCEKVKQK
CVKSCNCTELDQVENLHLDLCLAGNQLDLSKDSLGPTKSKIEGAGTSSISPPSP
ISPVASBSCTGLPLPLRPCEGSEMGKNSPENKWLMAAKRAENPSRSPSS
QTNSRRQSGKTLPSPVITPSSMRKICTYFHRKSQEDFCGPEHSTEL"

ORIGIN

		Query Match	98.5%; Score 2789.8; DB 5; Length 4221;
		Best Local Similarity	99.9%; Pred. No. 0;
		Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	22	CGATAACGATTTGTGTTGTGAGAGCGCAAGCTCGGATTTCTGCTGAACCTTGGAGGCATT	81
Db	1	CGATAACGATTTGTGTTGTGAGAGCGCAAGCTCGGATTTCTGCTGAACCTTGGAGGCATT	60
QY	82	TCTACGACTTTTCTCAGCTGAGGCTTTTCTCCGAGCCCTGATGCTCTTCAATTCGGTG	141
Db	61	TCTACGACTTTTCTCAGCTGAGGCTTTTCTCCGAGCCCTGATGCTCTTCAATTCGGTG	120
QY	142	CTCCGCGAGCCCGAGCTTGGCGTCTTGAGAAATGGATGGTCTTCAATACCCCTTCAA	201
Db	121	CTCCGCGAGCCCGAGCTTGGCGTCTTGAGAAATGGATGGTCTTCAATACCCCTTCAA	180
QY	202	TCCCTTCTGACTGGTTATCAGTCAGTCAGTGGTAAATGATGAACACACTTCTTATGGAGAAC	261
Db	181	TCCCTTCTGACTGGTTATCAGTCAGTGGTAAATGATGAACACACTTCTTATGGAGAAC	240
QY	262	GGAGTCCAGTTCTCTTTTGGATGTACCTTCTCTGCTGCCAATATGGAAACATGTA	321
Db	241	GGAGTCCAGTTCTCTTTTGGATGTACCTTCTCTGCTGCCAATATGGAAACATGTA	300
QY	322	CTAGCAGTTGCCAATGAAGAAAGGCTTTGTTGATGTATATAACAGAAATCAAAAGTTTC	381
Db	301	CTAGCAGTTGCCAATGAAGAAAGGCTTTGTTGATGTATATAACAGAAATCAAAAGTTTC	360
QY	382	AGAAAGAGTGCTTCAAGAGATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	441
Db	361	AGAAAGAGTGCTTCAAGAGATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	420
QY	442	GTTCTGCTGAACTTAAACTTTGTACAGCAGCAGTGATCAAAACAGCCAAATTTTGGGAC	501
Db	421	GTTCTGCTGAACTTAAACTTTGTACAGCAGCAGTGATCAAAACAGCCAAATTTTGGGAC	480
QY	502	GTAAGAGCTGCTGAGCTGATTTGGAACATGCAAGAGTCAATCAATGAGCCCTCAAGTCAGTT	561
Db	481	GTAAGAGCTGCTGAGCTGATTTGGAACATGCAAGAGTCAATCAATGAGCCCTCAAGTCAGTT	540
QY	562	GCTTTTCTTAAGTTTGAAAGCTGTATTCTGTACGGGTGGAAGATGGCAACATTATG	621
Db	541	GCTTTTCTTAAGTTTGAAAGCTGTATTCTGTACGGGTGGAAGATGGCAACATTATG	600
QY	622	GTCTGGGATACAGGTCAACAAAGATGGGTTTTATAGGCAAGTCAATCAATCAGT	681
Db	601	GTCTGGGATACAGGTCAACAAAGATGGGTTTTATAGGCAAGTCAATCAATCAGT	660
QY	682	GGAGCTCAATATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAGAAACAGAAATTC	741
Db	661	GGAGCTCAATATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAGAAACAGAAATTC	720

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 9, 2006, 05:50:06 ; Search time 1688 Seconds
(without alignments)
11693.386 Million cell updates/sec

Title: US-10-726-160-1
Perfect score: 2831
Sequence: 1 ggcacgagcgagggttgag.....atgaataaaaaaaaaaaaaa 2831

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 8:
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*
15: Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2831	100.0	2831	10	Adf69739 Gene enco
2	2831	100.0	2831	14	Adx69660 Human ret
3	2816.2	99.5	2838	4	Aah99153 Human pro
4	2804	99.0	4422	4	Aah72681 Human cer
5	2804	99.0	4422	5	Abv25710 Human pro
6	2804	99.0	4422	5	Adl61974 Human ova
7	2802	99.0	4450	10	Adf82011 Leukaemia
8	2789.8	98.5	4221	10	Aad63212 Human DNA
9	2789.8	98.5	4221	10	Aad62813 Human DNA
10	2789.8	98.5	4221	10	Abx95498 Human ret
11	2789.8	98.5	4221	11	Adp88308 Metastati
12	2789.8	98.5	4221	12	Adp07299 Human L2D
13	2789.8	98.5	4221	12	Ado20068 Human PRO
14	2789.8	98.5	4221	12	Ado20231 Human PRO
15	2789.8	98.5	4221	13	Adr25784 Breast ca
16	2789.8	98.5	4221	14	Ady17437 DNA encod
17	2789.8	98.5	4221	14	Ady61863 Human gen
18	2442.4	86.3	4083	12	Adq97696 Human can

19	2290.6	80.9	3685	4	AAH17860	Aah17860 Human CDN
20	1833	64.7	1833	4	AAH14095	Aah14095 Human CDN
21	1520	53.7	1871	5	ABA82997	Aba82997 Human tra
22	1302.2	46.0	2426	12	ADQ97693	Adq97693 Mouse can
23	1273.8	45.0	1291	4	AAH14025	Aah14025 Human CDN
24	1032.2	36.5	1111	5	ADM19613	Adm19613 Novel hum
25	856.4	30.3	905	4	AAH05270	Aah05270 Human CDN
26	835.8	29.5	8892	12	ADQ97695	Adq97695 Human can
27	804.4	28.4	853	4	AAH05411	Aah05411 Human CDN
28	799.2	28.2	1365	5	ADM19367	Adm19367 Novel hum
29	649.8	23.0	57860	12	ADQ97692	Adq97692 Mouse can
30	613	21.7	745	12	ADQ97698	Adq97698 Human can
31	558	19.7	708	2	AAH98989	Aah98989 Human val
32	542.8	19.2	760	4	AAH06624	Aah06624 Human CDN
33	541	19.1	541	10	ADF80608	Adf80608 Leukaemia
34	518.6	18.3	593	4	AAH10633	Aah10633 Human CDN
35	507	17.9	507	10	AAD59581	Aad59581 Human RA-
36	500.4	17.7	815	2	AAZ15581	Aaz15581 Human gen
37	451	15.9	451	14	ACL58353	ACL58353 Human col
38	440	15.5	468	9	ACH32745	ACH32745 Human end
39	439	15.5	543	5	ADL42840	Adl42840 Human ova
40	430.6	15.2	508	5	ADL44193	Adl44193 Human ova
41	421.4	14.9	445	4	AAH99154	Aah99154 83PSG4 su
42	421.4	14.9	477	2	AAZ17642	Aaz17642 Human gen
43	411.6	14.5	434	5	ADL37809	Adl37809 Human ova
44	411.6	14.5	434	5	ADI72670	Adi72670 Human ova
45	403.4	14.2	493	4	AAH33425	Aah33425 Human col

ALIGNMENTS

RESULT 1

ADf69739
ID ADF69739 standard; DNA; 2831 BP.

AC ADF69739;

DT 12-FEB-2004 (first entry)

XX Gene encoding human retinoic acid (RA) regulated protein, 8.31.

XX Human; retinoic acid regulated gene; retinoic acid regulated protein; RA;

KW Mitogen; morphogen; stem cell progenitor cell;

KW embryonic brain development; lung; liver; kidney;

KW cancer cell differentiation; Ushers Disease; cancer; cytostatic;

XX chromosome lq32.1-32.2; gene; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 124..2316

FT /tag= a

FT /product= "Retinoic acid (RA) regulated protein, 8.31"

XX US2003162268-A1.

XX 28-AUG-2003.

XX 08-APR-2003; 2003US-00409511.

XX 14-JUL-1999; 99US-00354359.

XX (IPNY/) IP N Y.

XX (CHEU/) W CHEUNG W M.

XX Ip NY, Cheung W M;

XX WPI; 2003-897935/82.

XX P-PSDB; ADF69740.

XX New retinoic acid regulated gene, useful in preparing a composition for treating or preventing Ushers Disease or cancer.

XX Claim 1; SEQ ID NO 1; 17pp; English.

XX The present invention relates to the isolation of a novel human retinoic acid (RA) regulated gene designated 8.31, and the polypeptide sequence it encodes. The gene encoding the 8.31 protein maps to human chromosome 1q32.1-32.2. The 8.31 protein is a mitogen or morphogen. It affects the identity of stem cell progenitor cells or the development of embryonic brain, lung, liver or kidney tissues. It also affects cancer cell differentiation. The 8.31 protein is useful in treating or preventing Ushers Disease or cancer. The present sequence encodes human retinoic acid (RA) regulated protein, 8.31.

XX Sequence 2831 BP; 807 A; 650 C; 607 G; 757 T; 0 U; 0 Other;

Query Match 100.0%; Score 2831; DB 10; Length 2831;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2831; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGAGGGAGTTGGAGGCGATAACGATTTGTGTGTGAGAGCGCAACGTGCGATT 60
Dy |||||
QY 1 GGCACGAGGGAGTTGGAGGCGATAACGATTTGTGTGTGAGAGCGCAACGTGCGATT 60
Dy |||||

QY 61 TCTGCTGAACCTTGGAGGCAATTTCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACC 120
Dy |||||

QY 61 TCTGCTGAACCTTGGAGGCAATTTCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACC 120
Dy |||||

QY 121 CTGATGCTCTTCAATTCGGTGTCTCGCCAGCCAGCTTGGCGTCTGAGAAATGGATGG 180
Dy |||||

QY 121 CTGATGCTCTTCAATTCGGTGTCTCGCCAGCCAGCTTGGCGTCTGAGAAATGGATGG 180
Dy |||||

QY 181 TCTTCACAATACCCCTTCAATCCCTTCTGACTGTTATCAGTGCAGTGGTAAATGATGAA 240
Dy |||||

QY 181 TCTTCACAATACCCCTTCAATCCCTTCTGACTGTTATCAGTGCAGTGGTAAATGATGAA 240
Dy |||||

QY 241 CACACTTCTTATGGAGAACAGAGTCCAGTTCCTCTCTTGGATGACCTTCTCTTCT 300
Dy |||||

QY 241 CACACTTCTTATGGAGAACAGAGTCCAGTTCCTCTCTTGGATGACCTTCTCTTCT 300
Dy |||||

QY 301 GCTCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGGCTTTTGTTCGATTGTAT 360
Dy |||||

QY 301 GCTCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGGCTTTTGTTCGATTGTAT 360
Dy |||||

QY 361 AACACAGAAATCAAAAGTTTCAAAAGAGTGCCTTCAAAAGATGAGTGGCTCACTGGAAT 420
Dy |||||

QY 361 AACACAGAAATCAAAAGTTTCAAAAGAGTGCCTTCAAAAGATGAGTGGCTCACTGGAAT 420
Dy |||||

QY 421 GCCGTCTTGTACCTGSCCTGGGTTCTGTGAACCTTAACTTTTACAGCAGCAGTGTAT 480
Dy |||||

QY 421 GCCGTCTTGTACCTGSCCTGGGTTCTGTGAACCTTAACTTTTACAGCAGCAGTGTAT 480
Dy |||||

QY 481 CAAACAGCCAAATTTTGGGACGTAAAGAGTGGTGAAGCTGATTGGAACATGCAAAAGTGCAT 540
Dy |||||

QY 481 CAAACAGCCAAATTTTGGGACGTAAAGAGTGGTGAAGCTGATTGGAACATGCAAAAGTGCAT 540
Dy |||||

QY 541 CAATGAGCCTCAAGTCAAGTTCCTTTCTAAGTTTGAAGCTGTATTTCTGTACGGGT 600
Dy |||||

QY 541 CAATGAGCCTCAAGTCAAGTTCCTTTCTAAGTTTGAAGCTGTATTTCTGTACGGGT 600
Dy |||||

QY 601 GGAAGAGATGGCAACATTTATGCTGTGGATACCGAGTGCACAAAAGATGGGTTTTAT 660
Dy |||||

QY 601 GGAAGAGATGGCAACATTTATGCTGTGGATACCGAGTGCACAAAAGATGGGTTTTAT 660
Dy |||||

QY 661 AGGCAAGTGAATCAAAATCAGTGGAGCTCACATACCTCAGACAAGCAACCCCTTCAAAA 720
Dy |||||

QY 661 AGGCAAGTGAATCAAAATCAGTGGAGCTCACATACCTCAGACAAGCAACCCCTTCAAAA 720
Dy |||||

QY 721 CCCAAGAGAAACAGAAATTCAAAAGGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTGT 780
Dy |||||

QY 721 CCCAAGAGAAACAGAAATTCAAAAGGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTGT 780
Dy |||||

QY 781 ACTGTGGTCTCTTTCAAGACAGAAATACCTTAGTCTCAGCAGGAGCTGTGATGGGATA 840
Dy |||||

781 ACTGTGGTCTCTTTCAAGACAGAGATACTTAGTCTCAGCAGGAGCTGTGATGGGATA 840
QY |||||

841 ATCAAAGTATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAGAACCCATAGCATCC 900
Dy |||||

841 ATCAAAGTATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAGAACCCATAGCATCC 900
Dy |||||

901 AAGTCTTTCTGTACCCAGGTAGCAGCACCTCGAAACCTTCGATATTCAAGTCTGATTTTG 960
QY |||||

901 AAGTCTTTCTGTACCCAGGTAGCAGCACCTCGAAACCTTCGATATTCAAGTCTGATTTTG 960
Dy |||||

961 GATTCCACTGGCTCTACTTTTATTTGCTAAATTGACACAGCATACATCTACATCTTTAAT 1020
QY |||||

961 GATTCCACTGGCTCTACTTTTATTTGCTAAATTGACACAGCATACATCTACATCTTTAAT 1020
Dy |||||

1021 ATGACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGAGACACCAAGAACTCTACCTTT 1080
QY |||||

1021 ATGACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGAGACACCAAGAACTCTACCTTT 1080
Dy |||||

1081 TATGTAATAATCCAGCCTTAGTCCAGATGACCAAGTTTATGTCAGTGGCTCAAGTATGATA 1140
QY |||||

1081 TATGTAATAATCCAGCCTTAGTCCAGATGACCAAGTTTATGTCAGTGGCTCAAGTATGATA 1140
Dy |||||

1141 GCTGCCTACATATGGAAGGCTCTCCACACCTCGCAACCTCTCTACTGTGCTCTGGGTCTAT 1200
QY |||||

1141 GCTGCCTACATATGGAAGGCTCTCCACACCTCGCAACCTCTCTACTGTGCTCTGGGTCTAT 1200
Dy |||||

1201 TCTCAAGAGGTACGTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY |||||

1201 TCTCAAGAGGTACGTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Dy |||||

1261 TCTGATGACAAATACACATAAATACTGGCGCTTGAATAGAGCTTTAGAGGAGAAACACGAGA 1320
QY |||||

1261 TCTGATGACAAATACACATAAATACTGGCGCTTGAATAGAGCTTTAGAGGAGAAACACGAGA 1320
Dy |||||

1321 GGTGATAAACTTTTCCAGGTTGGGCTCTCAGAAAGAAAAGAGTCAAGA CTTGGC 1380
QY |||||

1321 GGTGATAAACTTTTCCAGGTTGGGCTCTCAGAAAGAAAAGAGTCAAGA CTTGGC 1380
Dy |||||

1381 CTAGTAA CAGTAAGTAGCAGTAGCAGTAGTCTCTGCGCAAGCCCGCCAGGTTAAAGTGAAT 1440
QY |||||

1381 CTAGTAA CAGTAAGTAGCAGTAGCAGTAGTCTCTGCGCAAGCCCGCCAGGTTAAAGTGAAT 1440
Dy |||||

1441 CCATCCAAATTTTCCCGCTCATCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCT 1500
QY |||||

1441 CCATCCAAATTTTCCCGCTCATCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCT 1500
Dy |||||

1501 CTTCTTTCAAATACTCTCTAGCTTCTCTATTAAACCTCTCTGCAAGGCCCGGTCTCCC 1560
QY |||||

1501 CTTCTTTCAAATACTCTCTAGCTTCTCTATTAAACCTCTCTGCAAGGCCCGGTCTCCC 1560
Dy |||||

1561 ATCAACAGAGAGGCTCTGTCTCTCCGTCTCTCCAGGCCACCTTCTCTTCAAGATG 1620
QY |||||

1561 ATCAACAGAGAGGCTCTGTCTCTCCGTCTCTCCAGGCCACCTTCTCTTCAAGATG 1620
Dy |||||

1621 TCGATTAGAACTGGGTGACCCGAAACCTTCTCTCATCACCCCATCATCTCCACCTGCT 1680
QY |||||

1621 TCGATTAGAACTGGGTGACCCGAAACCTTCTCTCATCACCCCATCATCTCCACCTGCT 1680
Dy |||||

1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCAGAGTCACTCC 1740
QY |||||

1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCAGAGTCACTCC 1740
Dy |||||

1741 CAACGAGAGGCTCTCTGTAGCTTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGT 1800
QY |||||

1741 CAACGAGAGGCTCTCTGTAGCTTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGT 1800
Dy |||||

1801 CTGAGAGTGTGAACAAAGAGTGTGAAGTTGTAACTGTGTGACTGAGCTTTGATGGC 1860
QY |||||

1801 CTGAGAGTGTGAACAAAGAGTGTGAAGTTGTAACTGTGTGACTGAGCTTTGATGGC 1860
Dy |||||

1861 CAAAGTGAAGAAATCTTCAATTTGGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
QY |||||

1861 CAAAGTGAAGAAATCTTCAATTTGGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Dy |||||

QY 1921 AAGGACTCTTAGTCTCTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCA 1980
Db 1921 AAGGACTCTTAGTCTCTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCA 1980
QY 1981 GAGCCTCCGCTCTCTATCAGTCGCTATGCTTCAGAAAAGCTGTGGAAAGCTACCTCTTCT 2040
Db 1981 GAGCCTCCGCTCTCTATCAGTCGCTATGCTTCAGAAAAGCTGTGGAAAGCTACCTCTTCT 2040
QY 2041 TTGAGACCTTTGGAGAGGGCTCTGAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAT 2100
Db 2041 TTGAGACCTTTGGAGAGGGCTCTGAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAT 2100
QY 2101 AAAAACTGCTGTGTGGCCATGCGAGCAAAACGGAAGGCTGAGAATCCATCTCCACGAAGT 2160
Db 2101 AAAAACTGCTGTGTGGCCATGCGAGCAAAACGGAAGGCTGAGAATCCATCTCCACGAAGT 2160
QY 2161 CCGTCATCCAGACACCCCAATTCAGGAGACAGAGCGGAAGACATTTGCCAAGCCGGTTC 2220
Db 2161 CCGTCATCCAGACACCCCAATTCAGGAGACAGAGCGGAAGACATTTGCCAAGCCGGTTC 2220
QY 2221 ACCATCAGCCAGCTCCATGAGGAAAATCTGCACATACCTTCATGAGAAGTCCCGAGG 2280
Db 2221 ACCATCAGCCAGCTCCATGAGGAAAATCTGCACATACCTTCATGAGAAGTCCCGAGG 2280
QY 2281 GACTTCTGTGGTCTGAACTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTACTGA 2340
Db 2281 GACTTCTGTGGTCTGAACTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTACTGA 2340
QY 2341 GCTTTGGTCCACTAAAAAAGCTGAGCTTTGGTCCACTAAAAAAGAGTGAAGAAATACAA 2400
Db 2341 GCTTTGGTCCACTAAAAAAGCTGAGCTTTGGTCCACTAAAAAAGAGTGAAGAAATACAA 2400
QY 2401 AGTGACTCTAATCTGCTCTTTAAGAAAGCTGCTTTTCAATTTTGAACAAAATCTTT 2460
Db 2401 AGTGACTCTAATCTGCTCTTTAAGAAAGCTGCTTTTCAATTTTGAACAAAATCTTT 2460
QY 2461 TCAAGCTCAAAATGACCTAATCTGCTCTTACTACCAATATGATATGAGTATCTCCGAG 2520
Db 2461 TCAAGCTCAAAATGACCTAATCTGCTCTTACTACCAATATGATATGAGTATCTCCGAG 2520
QY 2521 GATGAATGCTGTGTTTAAATTTTCAATAAAGTAAATTTGTCACCTAGCATTTTGAATGAAT 2580
Db 2521 GATGAATGCTGTGTTTAAATTTTCAATAAAGTAAATTTGTCACCTAGCATTTTGAATGAAT 2580
QY 2581 AGCTTCACTTTTAAATTTTCACTCTCTATATAATGACATCCCAAGTTTCATGGAGG 2640
Db 2581 AGCTTCACTTTTAAATTTTCACTCTCTATATAATGACATCCCAAGTTTCATGGAGG 2640
QY 2641 CAATAAACAAGTTTCTGTTATCTCTGAACCTTTCTATGCTCAGTGAAGATATCTGCCAG 2700
Db 2641 CAATAAACAAGTTTCTGTTATCTCTGAACCTTTCTATGCTCAGTGAAGATATCTGCCAG 2700
QY 2701 CCACAGCATGAGGCTGTGAAGCTGACCTGAGAAATCTCTGCTGAAGACCTCTGTTCT 2760
Db 2701 CCACAGCATGAGGCTGTGAAGCTGACCTGAGAAATCTCTGCTGAAGACCTCTGTTCT 2760
QY 2761 GTTCTGCTCCCAACATGTATAATTTTATTTGAAATAATAATCTTTTCTATGAAAAA 2820
Db 2761 GTTCTGCTCCCAACATGTATAATTTTATTTGAAATAATAATCTTTTCTATGAAAAA 2820
QY 2821 AAAAAA 2831
Db 2821 AAAAAA 2831

RESULT 2
ADX69660
ID ADX69660 standard; cDNA; 2831 BP.
XX
AC ADX69660;
XX
DT 05-MAY-2005 (first entry)

XX Human retinoic acid regulated nuclear matrix protein (RAMP) cDNA.
DE DNA purification; retinoic acid; diagnosis; hepatocellular carcinoma;
XX cytosatic; prognosis; cancer; liver tumor;
KW retinoic acid regulated nuclear matrix protein; gene; ds.
XX Homo sapiens.
XX OS
XX Key Location/Qualifiers
FT 124..2316
CDS /*tag= a
FT /product= "retinoic acid regulated nuclear matrix
protein"
XX
XX US2005037372-A1.
XX 17-FEB-2005.
XX 02-DEC-2003; 2003US-00726160.
XX 14-JUL-1999; 99US-00354359.
XX 08-APR-2003; 2003US-00409511.
XX (UVHK-) UNIV HONG KONG SCI & TECHNOLOGY.
XX Ip NY, Cheung WNW;
PI WPI; 2005-172246/18.
XX P-PSDB; ADX69661.
DR Novel isolated retinoic acid regulated gene, useful for screening and
PT determining prognosis of patient having Hepatocellular cancer.
XX Claim 1; SEQ ID NO 1; 21bp; English.
XX The invention relates to an isolated retinoic acid regulated gene (I)
CC having a fully defined 2831 nucleotides sequence given in specification.
CC A protein (II) encoded by (I) is useful as a screening tool for
CC diagnosing Hepatocellular carcinomas and for monitoring treatment or
CC progression of Hepatocellular carcinomas. An antibody (III) to (II) is
CC useful for screening and determining the prognosis of a patient having
CC Hepatocellular cancer, where the biological samples comprise liver
CC tissues. This sequence corresponds to the cDNA encoding the retinoic acid
CC regulated nuclear matrix protein (RAMP; I).
XX Sequence 2831 BP; 807 A; 660 C; 607 G; 757 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 2831; DB 14; Length 2831;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2831; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGAGCGGAGTTGGAGCGATACGATTTGTTGTGTGAGAGCGCAACGTCGATT 60
Db 1 GGCACGAGCGGAGTTGGAGCGATACGATTTGTTGTGTGAGAGCGCAACGTCGATT 60
QY 61 TCTGCTGAACCTTGGAGGCAATTTACGACTTTTCTCAGCTGAGGCTTTCTCCGACC 120
Db 61 TCTGCTGAACCTTGGAGGCAATTTACGACTTTTCTCAGCTGAGGCTTTCTCCGACC 120
QY 121 CTGATGCTCTTCAATTCGGTGTCTCCGAGCCAGCTTGGCGTCTGAGAAATGGATGG 180
Db 121 CTGATGCTCTTCAATTCGGTGTCTCCGAGCCAGCTTGGCGTCTGAGAAATGGATGG 180
QY 181 TCTTCACAATACCTCTTCAATTCCTTCTGACTGGTATCAGTGCAGTGGTAAATGATGA 240
Db 181 TCTTCACAATACCTCTTCAATTCCTTCTGACTGGTATCAGTGCAGTGGTAAATGATGA 240
QY 241 CACACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTTTTGGATGATACCTTCTTCT 300
Db 241 CACACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTTTTGGATGATACCTTCTTCT 300
QY 301 GCTCCCAATATGGAACATGTACTAGCTTGCCAATGAAGGCTTTGTTTCGATTGTAT 360

||||| 301 GCTCCCAATATGGAACATGTACTAGAGTTGCCAATGAAGAGGCTTTGTCGATGTAAT 360
QY 361 AACACAGAAATCAAAAGTTTCAGAAAGAGTGTCTTCAAGAAATGATGGCTCACTGGAAAT 420
Db 361 AACACAGAAATCAAAAGTTTCAGAAAGAGTGTCTTCAAGAAATGATGGCTCACTGGAAAT 420
QY 421 GCGGCTTTTGACCTGGCCCTGGGTTCTGTTGTAACCTTAAACTTTGTTACAGCAGCAGGTGAT 480
Db 421 GCGGCTTTTGACCTGGCCCTGGGTTCTGTTGTAACCTTAAACTTTGTTACAGCAGCAGGTGAT 480
QY 481 CAAACAGCCAAATTTTGGGACCTAAAGCTGGTGAGCTGATTGGAAATGCAAAAGGTCAT 540
Db 481 CAAACAGCCAAATTTTGGGACCTAAAGCTGGTGAGCTGATTGGAAATGCAAAAGGTCAT 540
QY 541 CAATCAGCCTCAAGTCAAGTTCGCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGT 600
Db 541 CAATCAGCCTCAAGTCAAGTTCGCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGT 600
QY 601 GGAAGAGATGGCAACATTTATGTTCTGGGATACAGGTGCAAAAAGATGGGTTTTAT 660
Db 601 GGAAGAGATGGCAACATTTATGTTCTGGGATACAGGTGCAAAAAGATGGGTTTTAT 660
QY 661 AGGCAAGTGAATCAAAATCAGTGGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCAAAA 720
Db 661 AGGCAAGTGAATCAAAATCAGTGGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCAAAA 720
QY 721 CCCAAGAGAAACAGAAATTCAAAAGGACTTGCTCTCTGTGGATTTCCAGCAAAAGTGT 780
Db 721 CCCAAGAGAAACAGAAATTCAAAAGGACTTGCTCTCTGTGGATTTCCAGCAAAAGTGT 780
QY 781 ACTGTGGTCTCTTTCAAGACAGAAATACCTTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
Db 781 ACTGTGGTCTCTTTCAAGACAGAAATACCTTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
QY 841 ATCAAAGTATGGATTTACGTAAAGATTAATCTGTTATCGACAAGAACCCATGATCATCC 900
Db 841 ATCAAAGTATGGATTTACGTAAAGATTAATCTGTTATCGACAAGAACCCATGATCATCC 900
QY 901 AAGTCTTTCTGTATCCAGGTAGCAGCACTCGAAAACCTTGGATATTCAAAGTCTGATTTTG 960
Db 901 AAGTCTTTCTGTATCCAGGTAGCAGCACTCGAAAACCTTGGATATTCAAAGTCTGATTTTG 960
QY 961 GATTCACCTGGCTCTACTTTATTTGCTAAATTTGCAAGAGGATAATCATCTACATGTTTAAAT 1020
Db 961 GATTCACCTGGCTCTACTTTATTTGCTAAATTTGCAAGAGGATAATCATCTACATGTTTAAAT 1020
QY 1021 ATGACTGGGTTCAAGACTTCTCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTT 1080
Db 1021 ATGACTGGGTTCAAGACTTCTCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTT 1080
QY 1081 TATGTAAATCCAGCCTTAGTCCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAATGAA 1140
Db 1081 TATGTAAATCCAGCCTTAGTCCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAATGAA 1140
QY 1141 GCTGCTCATATGGAAGTCTCACAACCTGGGAAACCTCTCTACTGTGCTCTGCTGGGTGAT 1200
Db 1141 GCTGCTCATATGGAAGTCTCACAACCTGGGAAACCTCTCTACTGTGCTCTGCTGGGTGAT 1200
QY 1201 TCTCAAGAGGTCACGCTGTGTGCTGGTGTCTCACTCTGACTTCAAAAGATTTGCTACCTGT 1260
Db 1201 TCTCAAGAGGTCACGCTGTGTGCTGGTGTCTCACTCTGACTTCAAAAGATTTGCTACCTGT 1260
QY 1261 TCTGATGACAAATACATACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACCAAG 1320
Db 1261 TCTGATGACAAATACATACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACCAAG 1320
QY 1321 GGTGATAAACTTTTCCACGCTGGGTTGGGCTCTCAGAGAAAGATTCAGACCTGGC 1380
Db 1321 GGTGATAAACTTTTCCACGCTGGGTTGGGCTCTCAGAGAAAGATTCAGACCTGGC 1380
QY 1381 CTAGTAAACAGTAAACAGTACAGAGTACTCTCTGCCAAAGCCCCAGGGTAAAGTGAAT 1440

Db 1381 CTAGTAAACAGTAAACAGTAGCCAGAGTACTCTCTGCCAAAGCCCCAGGGTAAAGTGAAT 1440
QY 1441 CCATCCAAATTTTCCCGTCTATCCGAGCTTGTGCCCCCAAGCTGTGTGGAGACTTCCT 1500
Db 1441 CCATCCAAATTTTCCCGTCTATCCGAGCTTGTGCCCCCAAGCTGTGTGGAGACTTCCT 1500
QY 1501 CTTCTTTCAAAATCTCTAGCTTCTCTATTAAACCTCTCTCTGCAAGGCCGGTCTCTCC 1560
Db 1501 CTTCTTTCAAAATCTCTAGCTTCTCTATTAAACCTCTCTCTGCAAGGCCGGTCTCTCC 1560
QY 1561 ATCAACAGAGAGGCTCTGTCTCTCTCTCCAGCCACCTTCTCATCTTTCAAGATG 1620
Db 1561 ATCAACAGAGAGGCTCTGTCTCTCTCTCCAGCCACCTTCTCATCTTTCAAGATG 1620
QY 1621 TCGATTAGAAACTCGGGTGACCCGAACACCTTCTCTCATCACCCACCTCACTCCACCTGCT 1680
Db 1621 TCGATTAGAAACTCGGGTGACCCGAACACCTTCTCTCATCACCCACCTCACTCCACCTGCT 1680
QY 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCCTGTGAGCCAGAACTCATCC 1740
Db 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCCTGTGAGCCAGAACTCATCC 1740
QY 1741 CAAGCAGAGGCTTCTCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGT 1800
Db 1741 CAAGCAGAGGCTTCTCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGT 1800
QY 1801 CTGAGAGTGTGAAACAAAAGTGTGAAAGTTGTAACTGTGTGACTGAGCTTGTATGGC 1860
Db 1801 CTGAGAGTGTGAAACAAAAGTGTGAAAGTTGTAACTGTGTGACTGAGCTTGTATGGC 1860
QY 1861 CAAGTTCGAAATCTTCAATTTGGATCTGTCTGCTCTGCTGTAACACAGGAAGACTTGT 1920
Db 1861 CAAGTTCGAAATCTTCAATTTGGATCTGTCTGCTCTGCTGTAACACAGGAAGACTTGT 1920
QY 1921 AAGGACTCTCTAGTCTCTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCA 1980
Db 1921 AAGGACTCTCTAGTCTCTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCA 1980
QY 1981 GAGCTCTGCTCTCTATCAGTCTGATTCCTTCAAGAAAGCTGTGAAACGCTACCTCTTCT 2040
Db 1981 GAGCTCTGCTCTCTATCAGTCTGATTCCTTCAAGAAAGCTGTGAAACGCTACCTCTTCT 2040
QY 2041 TTGAGACCTCTGTGAGAGGCTCTGAAATGTTAGGCAAGAGAAATAGTTCCTCCAGAGAA 2100
Db 2041 TTGAGACCTCTGTGAGAGGCTCTGAAATGTTAGGCAAGAGAAATAGTTCCTCCAGAGAA 2100
QY 2101 AAAAATCTGTTGTTGGCCATGGCAGCCAAAACGGAAGCTGAGAAATCCATCTCCACGAAGT 2160
Db 2101 AAAAATCTGTTGTTGGCCATGGCAGCCAAAACGGAAGCTGAGAAATCCATCTCCACGAAGT 2160
QY 2161 CCGTCAATCCAGACACCCAAATTCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTC 2220
Db 2161 CCGTCAATCCAGACACCCAAATTCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTC 2220
QY 2221 ACCATCAGCCCGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCAGAG 2280
Db 2221 ACCATCAGCCCGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCAGAG 2280
QY 2281 GACTTCTGTGCTCTGAACTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTATCTGA 2340
Db 2281 GACTTCTGTGCTCTGAACTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTATCTGA 2340
QY 2341 GCTTTGGTCCATGAAACCAAGCTGAGCTTTGGTCCATTAACCAAGATGAAATATCAAG 2400
Db 2341 GCTTTGGTCCATGAAACCAAGCTGAGCTTTGGTCCATTAACCAAGATGAAATATCAAG 2400
QY 2401 AGTGACTCTATACTCTGGTCTTTAAGAAAGCTGCTTTTCAATTTTAGACAAAATCTTT 2460
Db 2401 AGTGACTCTATACTCTGGTCTTTAAGAAAGCTGCTTTTCAATTTTAGACAAAATCTTT 2460
QY 2461 TCAACGCTGAAATGATGCTAAATCTGGTTCTACTACCAATATGATATGACAGCTTCCGAG 2520
Db 2461 TCAACGCTGAAATGATGCTAAATCTGGTTCTACTACCAATATGATATGACAGCTTCCGAG 2520

QY 2521 GATGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGTCTACCTACATTTTGAATGAAT 2580
DB 2521 GATGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGTCTACCTACATTTTGAATGAAT 2580
QY 2581 AGTCTTCACTTTTAAATTTTCAATCTTCTCTATAATAATGACATCCCGAGTTTCATGGAGG 2640
DB 2581 AGTCTTCACTTTTAAATTTTCAATCTTCTCTATAATAATGACATCCCGAGTTTCATGGAGG 2640
QY 2641 CAAAAACAAGTTTCTGTATCTCTGAACTTCTATGCTCAGTGAAGATATCTGCCAG 2700
DB 2641 CAAAAACAAGTTTCTGTATCTCTGAACTTCTATGCTCAGTGAAGATATCTGCCAG 2700
QY 2701 CCACAGCATGAGGCTGTGAGGCTCACTGAGAAATCCTCTGCTGAAGACCCCTGTTCT 2760
DB 2701 CCACAGCATGAGGCTGTGAGGCTCACTGAGAAATCCTCTGCTGAAGACCCCTGTTCT 2760
QY 2761 GTTCTGCCCTCCACATGATATAATTTTATTGAAATAACATAATCTTTCTACTATGAAAAA 2820
DB 2761 GTTCTGCCCTCCACATGATATAATTTTATTGAAATAACATAATCTTTCTACTATGAAAAA 2820
QY 2821 AAAAAA 2831
DB 2821 AAAAAA 2831

RESULT 3
AAH99153
ID AAH99153 standard; cDNA; 2838 BP.
AC AAH99153;
XX
XX
DT 04-DEC-2001 (first entry)
XX
DE Human prostate-related gene 83P5G4 cDNA.
XX
XX 83P5G4; PCR primer; DNA adaptor; prostate; testis; tissue; cancer; ss;
KW tumour; kidney; brain; bone; ovary; breast; pancreas; uterus; colon;
KW lung; cytostatic; gene therapy; antibody therapy; ribozyme; liver;
KW single chain monoclonal antibody; serum; blood; urine; bladder; cervix;
KW rectum; stomach; human; chromosome 1q31-q32.
XX
OS Homo sapiens.
XX
XX WO200159115-A2.
PN
XX
PD 16-AUG-2001.
XX
PF 09-FEB-2001; 2001WO-US004426.
XX
PR 09-FEB-2000; 2000US-0181261P.
XX
PA (UROC-) UROGENESYS INC.
XX
XX Hubert RS, Afar DEH, Challita-Eid PM, Faris M, Levin B;
PI Mitchell SC, Jakobovits A;
XX
XX WPI; 2001-514669/56.
DR P-PSDB; ANM25224.
XX
XX An isolated 83P5G4-related protein useful as a diagnostic and/or
PT therapeutic agent in multiple cancers such as prostate, bladder and bone
PT cancer.
XX
PS Claim 1; Fig 2; 112pp; English.
XX
XX The nucleic acid sequences represent the 83P5G4 gene and the primers and
CC adaptors used to amplify 83P5G4 DNA. 83P5G4 exhibits prostate specific
CC expression in normal adult tissue, but it is also aberrantly expressed in
CC many cancers including tumours of the prostate, testis, bladder, kidney,
CC brain, bone, cervix, uterus, ovary, breast, pancreas, stomach, rectum,
CC liver, colon and lung. The 83P5G4 polynucleotide, its related protein and
CC also peptide fragments of the protein are therefore useful for diagnosing

CC and treating cancer. A vector comprising a polynucleotide which encodes a
CC single chain monoclonal antibody, that immunospecifically binds to an
CC 83P5G4-related protein, and a ribozyme capable of cleaving a
CC polynucleotide having the 83P5G4 coding sequence, are both useful in the
CC preparation of a composition for treating a patient with a cancer that
CC expresses 83P5G4. The sequences can be used in diagnostic methods to
CC monitor the level of 83P5G4 gene products in serum, blood, urine and
CC tissue and to thereby detect the presence of cancerous cells
XX
SQ Sequence 2838 BP; 810 A; 661 C; 609 G; 758 T; 0 U; 0 Other;
Query Match 99.5%; Score 2816.2; DB 4; Length 2838;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2821; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 3 CACGAGCGGAGTTGGAGCGGATTAACGATTTGTGTGTGAGAGCGGCAACGTGCGATTTC 62
DB 9 CAGTGGCGGAGTTGGAGCGGATTAACGATTTGTGTGTGAGAGCGGCAACGTGCGATTTC 68
QY 63 TGCTGAATCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 122
DB 69 TGCTGAATCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 128
QY 123 GATGCTCTTCAATTCGGTGTCTCCGACGCCAGCTTGGCGTCTCTGAGAAATGGATGTC 182
DB 129 GATGCTCTTCAATTCGGCGCTCCGCGCAGCTTGGCGTCTCTGAGAAATGGATGTC 188
QY 183 TTCACAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTATGATGAACA 242
DB 189 TTCACAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTATGATGAACA 248
QY 243 CACTTCTTATGAGAAACAGAGTCCAGATTCCTCTCTTTGGATGATCTTCTCTCTTCGC 302
DB 249 CACTTCTTATGAGAAACAGAGTCCAGTCCCTCTCTTTGGATGATCTTCTCTCTTCGC 308
QY 303 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTCGATTTGATAA 362
DB 309 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTCGATTTGATAA 368
QY 363 CACAGAATCACAAAGTTTTCAGAAAGAGTCTTCAAGAATGGATGGCTCATCGAATGC 422
DB 369 CACAGAATCACAAAGTTTTCAGAAAGAGTCTTCAAGAATGGATGGCTCATCGAATGC 428
QY 423 GGTCTTTCACCTGGCGCTGGCTTCTGCTGAACTTAACTTTTACAGCAGCAGGTGATCA 482
DB 429 GGTCTTTCACCTGGCGCTGGCTTCTGCTGAACTTAACTTTTACAGCAGCAGGTGATCA 488
QY 483 AACAGCCAAATTTTGGGACGTAAAGCTGGTGAAGCTGATTTGGAACATGCAAGGTGATCA 542
DB 489 AACAGCCAAATTTTGGGACGTAAAGCTGGTGAAGCTGATTTGGAACATGCAAGGTGATCA 548
QY 543 ATGAGCCTCAAGTCAGTTGCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGACGGTGG 602
DB 549 ATGAGCCTCAAGTCAGTTGCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGACGGTGG 608
QY 603 AAGAGATGCAACATTTATGCTCTGGATACAGGTGCAACAAAGATGGGTTTATAG 662
DB 609 AAGAGATGCAACATTTATGCTCTGGATACAGGTGCAACAAAGATGGGTTTATAG 668
QY 663 GCAAGTGAATCAAATCAGTGGAGCTCACAATACTCAGACAGCAAAACCCCTTCAAAACC 722
DB 669 GCAAGTGAATCAAATCAGTGGAGCTCACAATACTCAGACAGCAAAACCCCTTCAAAACC 728
QY 723 CAAAGAAACAGAAATTCAAAAGGACTTGTCTCTCTGTTGTTTCCAGCAAGGTTTAC 782
DB 729 CAAAGAAACAGAAATTCAAAAGGACTTGTCTCTCTGTTGTTTCCAGCAAGGTTTAC 788
QY 783 TGTGCTCTCTTCAAGACAGGATACCTTAGTCTCAGCAGAGCTGTGGATGGGATAT 842
DB 789 TGTGCTCTCTTCAAGACAGGATACCTTAGTCTCAGCAGAGCTGTGGATGGGATAT 848
QY 843 CAAAGTATGGGATTTACCTGAAGAATTTATCTGCTTATCGACAAAGAACCATAGCATCAA 902

849 CAAAGTATGGGATTACGTAAGAAATATATCTGCTTATCGACAAGAACCCATAGCATCCAA 908
903 GTCTTTCTCTGATACCCAGGTAGCAGACCTCGAAAACTTGGATATTCAGATCTGATTTTGGGA 962
909 GTCTTTCTCTGATACCCAGGTAGCAGACCTCGAAAACTTGGATATTCAGATCTGATTTTGGGA 968
963 TTCCCACTGGCTCTACTTTATTTGCTAAATTTGGACAGCATACATCTACATGTTTAAATAT 1022
969 TTCCCACTGGCTCTACTTTATTTGCTAAATTTGGACAGCATACATCTACATGTTTAAATAT 1028
1023 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTAA 1082
1029 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTAA 1088
1083 TGTAAAAATCCAGCCTTAGTCCAGATGACCAAGTTTTTATGTCAGTGGCTCAAGTGATGAAGC 1142
1089 TGTAAAAATCCAGCCTTAGTCCAGATGACCAAGTTTTTATGTCAGTGGCTCAAGTGATGAAGC 1148
1143 TGCCTACATATGGAAGGCTCTCCACACCTCTGGCAACCTCTACTGTGCTCCTGGGTCAATTC 1202
1149 TGCCTACATATGGAAGGCTCTCCACACCTCTGGCAACCTCTACTGTGCTCCTGGGTCAATTC 1208
1203 TCAAGAGGTCAGCTCTGTGTGCTGGTCCATCTGACTTCAAAAAGATTTGCTACCTGTTC 1262
1209 TCAAGAGGTCAGCTCTGTGTGCTGGTCCATCTGACTTCAAAAAGATTTGCTACCTGTTC 1268
1263 TGAATGACAAATACACTAAATAATCTGGCGTTGGAATAGAGGCTTTAGAGGAGAAACAGGAGG 1322
1269 TGAATGACAAATACACTAAATAATCTGGCGTTGGAATAGAGGCTTTAGAGGAGAAACAGGAGG 1328
1323 TGAATAAATCTTCCAGGTGGTGGTCTCTCAGAGAAAGAGTCAAGACCTTGGCCT 1382
1329 TGAATAAATCTTCCAGGTGGTGGTCTCTCAGAGAAAGAGTCAAGACCTTGGCCT 1388
1383 AGTAACAGTAACGAGTACGAGAGTACTCTGCCAAAGCCCCAGGGTAAAGTGCATATCC 1442
1389 AGTAACAGTACGAGTACGAGAGTACTCTGCCAAAGCCCCAGGGTAAAGTGCATATCC 1448
1443 ATCCAAATCTTCCCGTCAATCCGAGTGTGTCGCCCAAGCTGTGTGAGAGACCTCCCTCT 1502
1449 ATCCAAATCTTCCCGTCAATCCGAGTGTGTCGCCCAAGCTGTGTGAGAGACCTCCCTCT 1508
1503 TCCCTCAATACTCTAGCTCTCTATTAACACCTCTCTGCCAAGGCCGCTCTCCCAT 1562
1509 TCCCTCAATACTCTAGCTCTCTATTAACACCTCTCTGCCAAGGCCGCTCTCCCAT 1568
1563 CAACAGAGAGGCTCTGTCTCTCCGTCCTCCCAAGCCACCTTCACTCTTTCAAGATGTC 1622
1569 CAACAGAGAGGCTCTGTCTCTCCGTCCTCTCCCAAGCCACCTTCACTCTTTCAAGATGTC 1628
1623 GATTAGAAATCTGGGTGACCCGAACACCTTCTCATCAACCCATCACTCCACCTTGTTC 1682
1629 GATTAGAAATCTGGGTGACCCGAACACCTTCTCATCAACCCATCACTCCACCTTGTTC 1688
1683 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTGTGAGCCGAGAGTCAATCCCA 1742
1689 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTGTGAGCCGAGAGTCAATCCCA 1748
1743 AGCAGAGGCTCTGCTCTGAGTCTAGAAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT 1802
1749 AGCAGAGGCTCTGCTCTGAGTCTAGAAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT 1808
1803 GGAGAGTGTGAACAAAAAGTGTGTGAAGATTTGTAATCTGTGTGACTGAGCTTGATGGCCA 1862
1809 GGAGAGTGTGAACAAAAAGTGTGTGAAGATTTGTAATCTGTGTGACTGAGCTTGATGGCCA 1868
1863 AGTTGAAATCTTCAATTTGGATCTGTGCTGCTTCTGTTAAACACAGGAGGAGCCCTTAGTAA 1922
1869 AGTTGAAATCTTCAATTTGGATCTGTGCTGCTTCTGTTAAACACAGGAGGAGCCCTTAGTAA 1928
1923 GGACTCTCTAGGCTCTTACCAAAATCAAGCAAAAATTTGAAGAGGCTGTGTACAGTATCTCAGA 1982
1929 GGACTCTCTAGGCTCTTACCAAAATCAAGCAAAAATTTGAAGAGGCTGTGTACAGTATCTCAGA 1988

QY 1983 GCCTCCGCTCTCTATCAGTCCGTATCTTCAGAAAGCTGTGGAACGCTACCTCTTCTCTTT 2042
Db 1989 GCCTCCGCTCTCTATCAGTCCGTATCTTCAGAAAGCTGTGGAACGCTACCTCTTCTCTTT 2048
QY 2043 GAGACCTTTGGGAGAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAATAA 2102
Db 2049 GAGACCTTTGGGAGAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAATAA 2108
QY 2103 AAACCTGGTGTGTGGCCATGGCAGCCAAACCGGAAGGCTGAGAATCCATCTCTCCCAAGTTC 2162
Db 2109 AAACCTGGTGTGTGGCCATGGCAGCCAAACCGGAAGGCTGAGAATCCATCTCTCCCAAGTTC 2168
QY 2163 GTATCCAGACACCCAAATTTCCAGAGACAGAGGCGGAAAGACATGTCGAAGCCCGGTTCAC 2222
Db 2169 GTATCCAGACACCCAAATTTCCAGAGACAGAGGCGGAAAGACATGTCGAAGCCCGGTTCAC 2228
QY 2223 CATCAGCCCGAGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCCAAGGAGGA 2282
Db 2229 CATCAGCCCGAGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCCAAGGAGGA 2288
QY 2283 CTTCTGTGGTCTGAAACACACACAGAAATATAGATTTAATCTGAGTGAATCTTACTGAGC 2342
Db 2289 CTTCTGTGGTCTGAAACACACACAGAAATATAGATTTAATCTGAGTGAATCTTACTGAGC 2348
QY 2343 TTTGGTCCACTAAAACAGCTGAGCTTTGGTCCACTAAACAGATGAAATATACAGAG 2402
Db 2349 TTTGGTCCACTAAAACAGCTGAGCTTTGGTCCACTAAACAGATGAAATATACAGAG 2408
QY 2403 TGACTCTATAACTCTGTGCTTTTAAAGAAAGCTGCTTTTCAATTTTATAGACAAAATCTTTTC 2462
Db 2409 TGACTCTATAACTCTGTGCTTTTAAAGAAAGCTGCTTTTCAATTTTATAGACAAAATCTTTTC 2468
QY 2463 AACCTGAAATGTATACCTAAATCTGGTCTTACTACCAATATGTATATGATGAGCTTCCCGAGGA 2522
Db 2469 AACCTGAAATGTATACCTAAATCTGGTCTTACTACCAATATGTATATGATGAGCTTCCCGAGGA 2528
QY 2523 TGAATGCTGTGTTTAAATTTTCAATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAG 2582
Db 2529 TGAATGCTGTGTTTAAATTTTCAATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAG 2588
QY 2583 TCTTCACTTTTAAATTTTCACTCTCTATAATAATGACATCCAGTTCATCGAGGCA 2642
Db 2589 TCTTCACTTTTAAATTTTCACTCTCTATAATAATGACATCCAGTTCATCGAGGCA 2648
QY 2643 AAAAACAAGTTTCTGTATCTGTAAGACTTTCTATGCTCAGTGGAAGTATCTGCCAGCC 2702
Db 2649 AAAAACAAGTTTCTGTATCTGTAAGACTTTCTATGCTCAGTGGAAGTATCTGCCAGCC 2708
QY 2703 ACAGATGAGGCTGTGAGGCTGACTGAGAAATCTCTGCTGAGACCCCTGCTCTGT 2762
Db 2709 ACAGATGAGGCTGTGAGGCTGACTGAGAAATCTCTGCTGAGACCCCTGCTCTGT 2768
QY 2763 TCTCCCTCCCAACATGTATAATTTTATTTGAATAACATAATCTTTTCACTATGAAAAAAA 2822
Db 2769 TCTCCCTCCCAACATGTATAATTTTATTTGAATAACATAATCTTTTCACTATGAAAAAAA 2828
QY 2823 AAAAAAAAAA 2831
Db 2829 AAAAAAAAAA 2837

RESULT 4

AAH72681

ID AAH72681 standard; cDNA; 4422 BP.

XX AAH72681;

AC AAH72681;

XX 19-SEP-2001 (first entry)

XX Human cervical cancer marker nucleic acid 3955.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

KW

XX Homo sapiens.
OS WO200142467-A2.
PN
XX
XX
PD 14-JUN-2001.
PF
XX
XX 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-018931SP.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
PI WPI; 2001-375006/39.
XX
DR
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
PT
XX
XX Claim 1; Page 796; 1051pp; English.
XX
XX The invention relates to novel genes (AH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 4422 BP; 1285 A; 929 C; 880 G; 1311 T; 0 U; 17 Other;

Query Match
Best Local Similarity 99.0%; Score 2804; DB 4; Length 4422;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

3 CACGAGCGGAGTTGGAGCGATAACGATTGTGTGTGAGAGGCGCAACGTGCGATTTC 62
||
8 CAGTGGCGGAGTTGGAGCGATACGATTGTGTGTGAGAGGCGCAACGTGCGATTTC 67
||
63 TCCTGAATCTGAGGAGATTTCTACGACTTTTCTCAGCTGAGGCTTTTCCTCCGACCT 122
||
68 TGCTGAATCTGAGGAGATTTCTACGACTTTTCTCAGCTGAGGCTTTTCCTCCGACCT 127
||
123 GATGCTCTCAATTCGGTCTCGCGCGCCCGCAGCTTGGCGTCTCAGAGAAATGGATGTC 182
||
128 GATGCTCTCAATTCGGTCTCGCGCGCCCGCAGCTTGGCGTCTCAGAGAAATGGATGTC 187
||
183 TTCAATATACCTCTTCAATCTCTGACTGTTATCAGTCAGTGGTAAATGATGAACA 242
||
188 TTCAATATACCTCTTCAATCTCTGACTGTTATCAGTCAGTGGTAAATGATGAACA 247
||
243 CACTTCTTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTTGGATGTACCTCTCTCTGC 302
||
248 CACTTCTTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTTGGATGTACCTCTCTCTGC 307
||
303 TCCCAATATGGAACATGACTAGCAGTTCGCAATGAGAGGCTTTGTCGATTGTATTA 362
||
308 TCCCAATATGGAACATGACTAGCAGTTCGCAATGAGAGGCTTTGTCGATTGTATTA 367
||
363 CACAGATCACAAAGTTTCAGAAAGAGTCTTCAAGAAATGATGAGTGGTCACTGGAATGC 422
||
368 CACAGATCACAAAGTTTCAGAAAGAGTCTTCAAGAAATGATGAGTGGTCACTGGAATGC 427
||
423 CGTCTTTGACCTGGCTGGCTTCTGGTGAACCTTAACTTTTACAGCAGCAGGTGATCA 482
||
428 CGTCTTTGACCTGGCTGGCTTCTGGTGAACCTTAACTTTTACAGCAGCAGGTGATCA 487
||

QY 483 AACAGCCAAATTTTGGGACGTAAAGAGCTGGTGGAGCTGATTGGAACATGCAAGGTCATCA 542
||
DB 488 AACAGCCAAATTTTGGGACGTAAAGAGCTGGTGGAGCTGATTGGAACATGCAAGGTCATCA 547
||
QY 543 ATGACAGCCTCAAGTCAGTTGCTTTTCTAAAGTTTGAAGAAAGCTGTATTCTGTACGGGTGG 602
||
DB 548 ATGACAGCCTCAAGTCAGTTGCTTTTCTAAAGTTTGAAGAAAGCTGTATTCTGTACGGGTGG 607
||
QY 603 AAGAGATGGCAACATTAATGCTCTGGGATACACAGGTGCAACAAAAGATGGGTTTATAG 662
||
DB 608 AAGAGATGGCAACATTAATGCTCTGGGATACACAGGTGCAACAAAAGATGGGTTTATAG 667
||
QY 663 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTTCAGACAAAGCAAAACCTTCAAAACC 722
||
DB 668 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTTCAGACAAAGCAAAACCTTCAAAACC 727
||
QY 723 CAAGAAGAAAACAGAAATTCAAAAGGACTTGTCTCTGTGGAATTTCCAGCAAGGTTTAC 782
||
DB 728 CAAGAAGAAAACAGAAATTCAAAAGGACTTGTCTCTGTGGAATTTCCAGCAAGGTTTAC 787
||
QY 783 TGTGCTCTCTTTCAGACGAGAAATACCTTAGTCTCTCAGCAGGAGCTGTGGATGGGATTAAT 842
||
DB 788 TGTGCTCTCTTTCAGACGAGAAATACCTTAGTCTCTCAGCAGGAGCTGTGGATGGGATTAAT 847
||
QY 843 CAAAGTATGSGATTTTACGTAAGAAATTTACTGCTTTATCGACAAAGAACCCATAGCATCCAA 902
||
DB 848 CAAAGTATGSGATTTTACGTAAGAAATTTACTGCTTTATCGACAAAGAACCCATAGCATCCAA 907
||
QY 903 GTCTTTCTCTGATCCAGGTAGCAGCAGTCTGAAAACTTTGGATATTCAAGTCTGATTTTGA 962
||
DB 908 GTCTTTCTCTGATCCAGGTAGCAGCAGTCTGAAAACTTTGGATATTCAAGTCTGATTTTGA 967
||
QY 963 TTCACCTGGCTCTACTTTATTGTTGTAATGACAGACGATTAACATCTACATGTTTAAATAT 1022
||
DB 968 TTCCACTGGCTCTACTTTATTGTTGTAATGACAGACGATTAACATCTACATGTTTAAATAT 1027
||
QY 1023 GACTGGGTTGAAGACTTCTCCAGTGGCTATTCTCAATGGACACACCAAGAACTCTACCTTTTA 1082
||
DB 1028 GACTGGGTTGAAGACTTCTCCAGTGGCTATTCTCAATGGACACACCAAGAACTCTACCTTTTA 1087
||
QY 1083 TGTAAATCCAGCCTTAGTCCAGATGACAGTCTTGTAGTGGCTCAAGTGATGAAGC 1142
||
DB 1088 TGTAAATCCAGCCTTAGTCCAGATGACAGTCTTGTAGTGGCTCAAGTGATGAAGC 1147
||
QY 1143 TGCCTACATATGGAAGGTCTCCACACCTGSCAACCTCTACTGTGCTCCTGGGTCAATTC 1202
||
DB 1148 TGCCTACATATGGAAGGTCTCCACACCTGSCAACCTCTACTGTGCTCCTGGGTCAATTC 1207
||
QY 1203 TCAGAGGTACAGTCTGTGCTGTGCTGTCATCTGACCTTCAAAAGATTGCTTACCTGTTC 1262
||
DB 1208 TCAGAGGTACAGTCTGTGCTGTGCTGTCATCTGACCTTCAAAAGATTGCTTACCTGTTC 1267
||
QY 1263 TGATGACAAATACACTTAAATAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGGAG 1322
||
DB 1268 TGATGACAAATACACTTAAATAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGGAG 1327
||
QY 1323 TGATAAACTTTTCCAGGTTGGGCTCTCAGAGAAAAAGAGTCAAGACCTTGGCT 1382
||
DB 1328 TGATAAACTTTTCCAGGTTGGGCTCTCAGAGAAAAAGAGTCAAGACCTTGGCT 1387
||
QY 1383 AGTAACAGTAAACAGTAGCCAGAGTACTCTCTGCCAAAGCCCCAGGGTAAAGTGCAATCC 1442
||
DB 1388 AGTAACAGTAAACAGTAGCCAGAGTACTCTCTGCCAAAGCCCCAGGGTAAAGTGCAATCC 1447
||
QY 1443 ATCCAAATTTTCCCGCTCATCCGAGCTTGTGCCCAAGCTGTCTGGAGACCTTCCCTCT 1502
||
DB 1448 ATCCAAATTTTCCCGCTCATCCGAGCTTGTGCCCAAGCTGTCTGGAGACCTTCCCTCT 1507
||
QY 1503 TCCTTCAAAATCTCCTTAGTCTCTATTTAAACCTCTCTCTGCCAAAGCCCGGTCTCCCAT 1562
||
DB 1508 TCCTTCAAAATCTCCTTAGTCTCTATTTAAACCTCTCTCTGCCAAAGCCCGGTCTCCCAT 1567
||

Db 8 CAGTGGCGGAGCTGGAGGCGATACGATTTGTGTGTGAGAGGCGCAAGCTGCGATTTC 67
Qy 63 TGCCTGAATTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCT 122
Db 68 TGCCTGAATTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCT 127
Qy 123 GATGCTCTTCAATTCGGTGTCTCGGCGAGCCGAGCTTTGGCGTCTTGAGAAATGGAATGGTC 182
Db 128 GATGCTCTTCAATTCGGTGTCTCGGCGAGCCGAGCTTTGGCGTCTTGAGAAATGGAATGGTC 187
Qy 183 TTCACAATACCTCTTCAATCCCTCTGACTGCTTATCAGTCAGTGGTGAATGATGAACA 242
Db 188 TTCACAATACCTCTTCAATCCCTCTGACTGCTTATCAGTCAGTGGTGAATGATGAACA 247
Qy 243 CACTCTTATGAGAAACAGGAGTCCGAGTCTCTCTTTTGGATGTACCTCTCTCTGCG 302
Db 248 CACTCTTATGAGAAACAGGAGTCCGAGTCTCTCTTTTGGATGTACCTCTCTCTGCG 307
Qy 303 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTTGTTCGATTGTATAA 362
Db 308 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTTGTTCGATTGTATAA 367
Qy 363 CACAGAAATCAAAAGTTTCAGAAAGAGTGTTCAAAGAAATGATGGTCACTGGAATGC 422
Db 368 CACAGAAATCAAAAGTTTCAGAAAGAGTGTTCAAAGAAATGATGGTCACTGGAATGC 427
Qy 423 CGTCTTTGACCTGGCTGGGTTCTGGTGAACTTAAACTTGTATCAGCAGCAGGTGATCA 482
Db 428 CGTCTTTGACCTGGCTGGGTTCTGGTGAACTTAAACTTGTATCAGCAGCAGGTGATCA 487
Qy 483 AACAGCCAAATTTTGGGAGCTAAAGCTGTGAGCTGATTGGAAATGCAAAAGGTCAATCA 542
Db 488 AACAGCCAAATTTTGGGAGCTAAAGCTGTGAGCTGATTGGAAATGCAAAAGGTCAATCA 547
Qy 543 ATGACGCTCAAGTCAAGTTCCTTTTCTAAAGTTTGAAGAACTGTATTCTGACGGGTGG 602
Db 548 ATGACGCTCAAGTCAAGTTCCTTTTCTAAAGTTTGAAGAACTGTATTCTGACGGGTGG 607
Qy 603 AAGAGTGGCAACATTTATGGTCTGGGATACAGGTGCAACAAAGATGGGTTTTATAG 662
Db 608 AAGAGTGGCAACATTTATGGTCTGGGATACAGGTGCAACAAAGATGGGTTTTATAG 667
Qy 663 GCAAGTGAATCAATCAGTGGAGCTCACAAATACCTCAGACAAGCAACCCCTTCAAAACC 722
Db 668 GCAAGTGAATCAATCAGTGGAGCTCACAAATACCTCAGACAAGCAACCCCTTCAAAACC 727
Qy 723 CAAGAAGAAACAGAAATTCAAAAGGACTGTCTCTCTGTTGATTTCCAGCAAGGTGTAC 782
Db 728 CAAGAAGAAACAGAAATTCAAAAGGACTGTCTCTCTGTTGATTTCCAGCAAGGTGTAC 787
Qy 783 TGTGCTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAT 842
Db 788 TGTGCTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAT 847
Qy 843 CAAAGTATGGGATTTACGTAAAGATTTATCTGCTTATCGACAAGAACCCATAGCATCCAA 902
Db 848 CAAAGTATGGGATTTACGTAAAGATTTATCTGCTTATCGACAAGAACCCATAGCATCCAA 907
Qy 903 GTCTTTCTGTATCCAGGTAGCAGCACTCGAAAACCTTGGATATTCAGGTCTGATTTTGA 962
Db 908 GTCTTTCTGTATCCAGGTAGCAGCACTCGAAAACCTTGGATATTCAGGTCTGATTTTGA 967
Qy 963 TTCCACTGGCTCTATTTTATTTGCTAAATTTGACAGACGATATCATCTGTTAATAT 1022
Db 968 TTCCACTGGCTCTATTTTATTTGCTAAATTTGACAGACGATATCATCTGTTAATAT 1027
Qy 1023 GACTGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGACACCGAGACTCTACCTTTTA 1082
Db 1028 GACTGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGACACCGAGACTCTACCTTTTA 1087
Qy 1083 TGTAAATCCAGCTTAGTCCAGATGACCAAGTTTTTTAGTCAAGTGGCTCAAGTGAAGC 1142
Db 1088 TGTAAATCCAGCTTAGTCCAGATGACCAAGTTTTTTAGTCAAGTGGCTCAAGTGAAGC 1147

Qy 1143 TGCCTACATATGGAAGTCTCCACACCTCGCAACCTCTACTGTGTCTCTGGGTCAATC 1202
Db 1148 TGCCTACATATGGAAGTCTCCACACCTCGCAACCTCTACTGTGTCTCTGGGTCAATC 1207
Qy 1203 TCAAGAGTCAAGTCTGTGTGTCTGGTGTCCATCTGACTTCAAAAGATTTGCTGATCT 1262
Db 1208 TCAAGAGTCAAGTCTGTGTGTCTGGTGTCCATCTGACTTCAAAAGATTTGCTGATCT 1267
Qy 1263 TGATGACATACATCAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAAACAGGAGG 1322
Db 1268 TGATGACATACATCAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAAACAGGAGG 1327
Qy 1323 TGATAAACTTTTCCAGGTGGGTGGGCTCTCAGAGAAAAAGAGTCAAGACCTTGGCT 1382
Db 1328 TGATAAACTTTTCCAGGTGGGTGGGCTCTCAGAGAAAAAGAGTCAAGACCTTGGCT 1387
Qy 1383 AGTAAACAGTAAACGAGTACGAGATCTCTGCGCAAAAGCCCGAGGGTAAAGTGAATCC 1442
Db 1388 AGTAAACAGTAAACGAGTACGAGATCTCTGCGCAAAAGCCCGAGGGTAAAGTGAATCC 1447
Qy 1443 ATCCAAATTTTCCCGCTCATCCGAGCTTGTGCGCAAGCTGTCTGGAGACCTCCCTCT 1502
Db 1448 ATCCAAATTTTCCCGCTCATCCGAGCTTGTGCGCAAGCTGTCTGGAGACCTCCCTCT 1507
Qy 1503 TCCTTTCAAAATCTCTACGTTCTCTATTAAAACTCTCTGCGCAAGGCCCGGTCTCCCAT 1562
Db 1508 TCCTTTCAAAATCTCTACGTTCTCTATTAAAACTCTCTGCGCAAGGCCCGGTCTCCCAT 1567
Qy 1563 CAACAGAGAGGCTCTGTCTCTCTCTCCCAAGCCACCTTCATCTTTCAAGATGC 1622
Db 1568 CAACAGAGAGGCTCTGTCTCTCTCTCCCAAGCCACCTTCATCTTTCAAGATGC 1627
Qy 1623 GATTAGAACTTGGGTGAGCCCGAAACACCTTCTCTCATCACCACCTCATCCCTGCTTC 1682
Db 1628 GATTAGAACTTGGGTGAGCCCGAAACACCTTCTCTCATCACCACCTCATCCCTGCTTC 1687
Qy 1683 GGAGACCAAGATCATGTCTCGAGAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCA 1742
Db 1688 GGAGACCAAGATCATGTCTCGAGAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCA 1747
Qy 1743 AGCAGAGCTTGTCTGAGTCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCT 1802
Db 1748 AGCAGAGCTTGTCTGAGTCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCT 1807
Qy 1803 GGAGAGTGTGAAAACAAAGTGTGAGAGTTGTAACTGTGTGACTGAGCTTGTATGSCCA 1862
Db 1808 GGAGAGTGTGAAAACAAAGTGTGAGAGTTGTAACTGTGTGACTGAGCTTGTATGSCCA 1867
Qy 1863 AGTTGMAAATCTTCAATTTGGATCTGTGCTGCTTGTGTAACCCAGGAAGACCTTAGTAA 1922
Db 1868 AGTTGMAAATCTTCAATTTGGATCTGTGCTGCTTGTGTAACCCAGGAAGACCTTAGTAA 1927
Qy 1923 GGACTCTCTAGGTCTTACCAAAATCAAGCAAAATTTGAAGAGCTGGTACCACTCTCAGA 1982
Db 1928 GGACTCTCTAGGTCTTACCAAAATCAAGCAAAATTTGAAGAGCTGGTACCACTCTCAGA 1987
Qy 1983 GCCTCCGTCTCTATCAGTCCGTATGCTTCAAGAAAGCTGTGGAAGCTTACCTCTCTCTTT 2042
Db 1988 GCCTCCGTCTCTATCAGTCCGTATGCTTCAAGAAAGCTGTGGAAGCTTACCTCTCTCTTT 2047
Qy 2043 GAGACCTTGTGAGAGGCTCTGAAATGTTAGGCAAAAGAGAAATAGTTCCTCCAGAGAAATA 2102
Db 2048 GAGACCTTGTGAGAGGCTCTGAAATGTTAGGCAAAAGAGAAATAGTTCCTCCAGAGAAATA 2107
Qy 2103 AAACCTGGTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCC 2162
Db 2108 AAACCTGGTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCC 2167
Qy 2163 GTCTATCCAGACACCCAAATTCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCA 2222
Db 2168 GTCTATCCAGACACCCAAATTCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCA 2227

2223	Qy	CATCACCCGACGCTCCATGAGGAAATCTGCA	CATCTTCCATAGAAAGTCCCGAGGGA	2282
2228	Db	CATCACCCGACGCTCCATGAGGAAATCTGCA	CATCTTCCATAGAAAGTCCCGAGGGA	2287
2283	Qy	CTTCTGGTGCCTGAAACACCTCAACAGAA	TTATAGATTCTAATCTGAGTGGAGTTACTGAGC	2342
2288	Db	CTTCTGGTGCCTGAAACACCTCAACAGAA	TTATAGATTCTAATCTGAGTGGAGTTACTGAGC	2347
2343	Qy	TTTGGTCCACTAAAAAAGCTGAGCTTTGGT	GCACCTAAAAACAAGATGAAAAATACAAAG	2402
2348	Db	TTTGGTCCACTAAAAAAGCTGAGCTTTGGT	GCACCTAAAAACAAGATGAAAAATACAAAG	2407
2403	Qy	TGACTCTATAACTCTGCTCTTTAAGAAAG	CTGCTCTTTTTCATTTTATAGACAAAATCTTTTC	2462
2408	Db	TGACTCTATAACTCTGCTCTTTAAGAAAG	CTGCTCTTTTTCATTTTATAGACAAAATCTTTTC	2467
2463	Qy	AACGCTGAAATGTACCTAATCTGGTTCTA	CTACCACTAATGTATATGACGCTTCCCGAGGA	2522
2468	Db	AACGCTGAAATGTACCTAATCTGGTTCTA	CTACCACTAATGTATATGACGCTTCCCGAGGA	2527
2523	Qy	TGAATGCTGTGTTTAAATTTTCAATAAG	TAAATTTGTCACTCTAGCATTTTGAATGAATAG	2582
2528	Db	TGAATGCTGTGTTTAAATTTTCAATAAG	TAAATTTGTCACTCTAGCATTTTGAATGAATAG	2587
2583	Qy	TCCTTCACTTTTTAAATTTATTCATCTTCT	TATATAATGACATCCCAGTTTCATGAGGCA	2642
2588	Db	TCCTTCACTTTTTAAATTTATTCATCTTCT	TATATAATGACATCCCAGTTTCATGAGGCA	2647
2643	Qy	AAAAACAAGTTTCTGTGTATCCGTGAAAC	TTTCTATGCTCAGTGGAAGTATCTGCCAGCC	2702
2648	Db	AAAAACAAGTTTCTGTGTATCCGTGAAAC	TTTCTATGCTCAGTGGAAGTATCTGCCAGCC	2707
2703	Qy	ACAGCATGAGGCCCTGTGAAGCGCTGAC	TGAGAAATCCCTCTGCTGAAGACCCCTGGTTCTGT	2762
2708	Db	ACAGCATGAGGCCCTGTGAAGCGCTGAC	TGAGAAATCCCTCTGCTGAAGACCCCTGGTTCTGT	2767
2763	Qy	TCCTGCCCTCCAAACATGTATAATTTTAT	TGTGAAATACATAAATCTTTTCTACTATG	2814
2768	Db	TCCTGCCCTCCAAACATGTATAATTTTAT	TGTGAAATACATAAATCTTTTCTACTATG	2819

RESULT 6
ADL61974
ID ADL61974 standard; DNA; 4422 BP.
XX
XX ADL61974;
XX
XX
XX DT (first entry)
XX
DE Human ovarian cancer DNA marker #20186.
XX
XX
KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.

XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
XX
PS Disclosure; SEQ ID NO 20186; 106pp; English.
XX
XX The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.

XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX
PI Lee J, Lillie J;
XX
-DR WPI; 2001-611502/70.

QY 423 CGTCTTTGACCTGGCGTTCCTGGTGAACTTAAACTTTGTACAGCAGCAGGTGATCA 482
Db 428 CGTCTTTGACCTGGCGTTCCTGGTGAACTTAAACTTTGTACAGCAGCAGGTGATCA 487
QY 483 AACAGCAAATTTTGGGACGTAAAGCTGGTGAGCTGATGGAAACATGCAAAAGGTGATCA 542
Db 488 AACAGCAAATTTTGGGACGTAAAGCTGGTGAGCTGATGGAAACATGCAAAAGGTGATCA 547
QY 543 ATGACGCTCAAGTCAAGTTCCTTTCTAAAGTTTGAGAAAGCTGTATTTCTGTACGGGTGG 602
Db 548 ATGACGCTCAAGTCAAGTTCCTTTCTAAAGTTTGAGAAAGCTGTATTTCTGTACGGGTGG 607
QY 603 AAGAGATGGCAACATTATGCTGCGGATACCAAGGTGCAACAAAAGATGGGTTTTATAG 662
Db 608 AAGAGATGGCAACATTATGCTGCGGATACCAAGGTGCAACAAAAGATGGGTTTTATAG 667
QY 663 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTCAGCAAGCAAAACCCCTTCAAAACC 722
Db 668 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTCAGCAAGCAAAACCCCTTCAAAACC 727
QY 723 CAAGAGAAACAGAAATTCAAAAGGACTTGCCTCTCTGTGGATTTCCAGCAAGGTGTAC 782
Db 728 CAAGAGAAACAGAAATTCAAAAGGACTTGCCTCTCTGTGGATTTCCAGCAAGGTGTAC 787
QY 783 TGTGCTCTCTTTCAAGACAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAAT 842
Db 788 TGTGCTCTCTTTCAAGACAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAAT 847
QY 843 CAAAGTATGGGATTTTACGTAAGAAATATATCTGCTTATCGACAAAGAACCCATAGCATCCAA 902
Db 848 CAAAGTATGGGATTTTACGTAAGAAATATATCTGCTTATCGACAAAGAACCCATAGCATCCAA 907
QY 903 GTCTTCTCGTACCGAGTACGAGCTCGAAGAACTTGGATATTCAGTCTGATTTTGA 962
Db 908 GTCTTCTCGTACCGAGTACGAGCTCGAAGAACTTGGATATTCAGTCTGATTTTGA 967
QY 963 TTCCACTGGCTCTACTTTATTTGCTAAATGACAGACGATAACATCTACATGTTTAAATAT 1022
Db 968 TTCCACTGGCTCTACTTTATTTGCTAAATGACAGACGATAACATCTACATGTTTAAATAT 1027
QY 1023 GACTGGTTGAAGACTTCTCAGTGGCTATTTTCAATGGACACCGAACTCTACCTTTTA 1082
Db 1028 GACTGGTTGAAGACTTCTCAGTGGCTATTTTCAATGGACACCGAACTCTACCTTTTA 1087
QY 1083 TGTAAATCCAGCTTAGTCCAGATGACAGCTTTTGTAGTGGCTCAAGTGATGAAGC 1142
Db 1088 TGTAAATCCAGCTTAGTCCAGATGACAGCTTTTGTAGTGGCTCAAGTGATGAAGC 1147
QY 1143 TGCCTACATATGGAAGGTCTCCACACCTGGCAACCTCTACTGTGCTCCTGGGTCAATTC 1202
Db 1148 TGCCTACATATGGAAGGTCTCCACACCTGGCAACCTCTACTGTGCTCCTGGGTCAATTC 1207
QY 1203 TCAAGAGGTCAAGTCTGTGCTGGTGTCCATCTGACTTCAAAAAGTTGCTACCTGTTTC 1262
Db 1208 TCAAGAGGTCAAGTCTGTGCTGGTGTCCATCTGACTTCAAAAAGTTGCTACCTGTTTC 1267
QY 1263 TGAATGACAACTACATAAAATCTGGGCTTGAATAGAGCTTAGAGGAAACACAGGAGG 1322
Db 1268 TGAATGACAACTACATAAAATCTGGGCTTGAATAGAGCTTAGAGGAAACACAGGAGG 1327
QY 1323 TGAATAACTTTCCACGGTGGGTTGGGCTCTCAGAGAAAGAGTCAAGACCTGGCT 1382
Db 1328 TGAATAACTTTCCACGGTGGGTTGGGCTCTCAGAGAAAGAGTCAAGACCTGGCT 1387
QY 1383 AGTAACAGTAAACAGTAGCAGAGTACTCTGCAAGAGCCCGGAGGTAAAGTCAATCC 1442
Db 1388 AGTAACAGTAAACAGTAGCAGAGTACTCTGCAAGAGCCCGGAGGTAAAGTCAATCC 1447
QY 1443 ATCCAAATCTTCCCGTCAATCCGAGCTTGTGCCCCAGAGCTGTGAGAGACCTCCCTCT 1502
Db 1448 ATCCAAATCTTCCCGTCAATCCGAGCTTGTGCCCCAGAGCTGTGAGAGACCTCCCTCT 1507
QY 1503 TCCTTCAAACTACTCCTACGTTCTCTATTAAACCTCTCTCTGCAAGGCCGGTCTCCCAT 1562

Db 1508 TCTTCAAACTACTCCTACGTTCTCTATTAAACCTCTCTCTGCAAGGCCGGTCTCCCAT 1567
QY 1563 CAACAGAGAGGCTCTGTCTCTCGTCTCTCCAGGCCACTTCACTTTCAAGATGTC 1622
Db 1568 CAACAGAGAGGCTCTGTCTCTCGTCTCTCCAGGCCACTTCACTTTCAAGATGTC 1627
QY 1623 GATTAGAACTCGGTGACCCGAAACACTTCTCATCACACCCCATCACTCCACCTGCTTC 1682
Db 1628 GATTAGAACTCGGTGACCCGAAACACTTCTCATCACACCCCATCACTCCACCTGCTTC 1687
QY 1683 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCAA 1742
Db 1688 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCAA 1747
QY 1743 AGCAGAGGCTTGGCTGTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT 1802
Db 1748 AGCAGAGGCTTGGCTGTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT 1807
QY 1803 GGAGAGTGTGAAAACAAAAGTGTGAAAGAGTGTAACTGTGTGACTGAGCTTGTATGGCCA 1862
Db 1808 GGAGAGTGTGAAAACAAAAGTGTGAAAGAGTGTAACTGTGTGACTGAGCTTGTATGGCCA 1867
QY 1863 AGTTGAAAATCTTCAATTTGGATCTGTGCTGCTTGTGTTAAACAGGAAGACCTTAGTAA 1922
Db 1868 AGTTGAAAATCTTCAATTTGGATCTGTGCTGCTTGTGTTAAACAGGAAGACCTTAGTAA 1927
QY 1923 GGACTCTCTAGGCTCTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACAGTATCTAGA 1982
Db 1928 GGACTCTCTAGGCTCTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACAGTATCTAGA 1987
QY 1983 GCCTCCGCTCTCCTATCAGTCCGTATGTTTCAAGAAAGCTGTGGAAGCTTACCTTCTCTTT 2042
Db 1988 GCCTCCGCTCTCCTATCAGTCCGTATGTTTCAAGAAAGCTGTGGAAGCTTACCTTCTCTTT 2047
QY 2043 GAGACCTTGTGGAGAGGCTCTGAAATGGTGAAGCAAGAGAAATAGTTCCCCAGAGAAATA 2102
Db 2048 GAGACCTTGTGGAGAGGCTCTGAAATGGTGAAGCAAGAGAAATAGTTCCCCAGAGAAATA 2107
QY 2103 AAACCTGGTGTGGGCTAGGAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAGTCC 2162
Db 2108 AAACCTGGTGTGGGCTAGGAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAGTCC 2167
QY 2163 GTCAATCCAGACACCCCAATTCAGAGACAGCGGAAAGACATTTGCCAGCCGGTCAAC 2222
Db 2168 GTCAATCCAGACACCCCAATTCAGAGACAGCGGAAAGACATTTGCCAGCCGGTCAAC 2227
QY 2223 CATCACGCCAGCTCCATGAGGAAATCTGCACATACCTTCCATAGAAAAGTCCCGAGGAGA 2282
Db 2228 CATCACGCCAGCTCCATGAGGAAATCTGCACATACCTTCCATAGAAAAGTCCCGAGGAGA 2287
QY 2283 CTTCTGTGGTCTGAAACACTCAACAGAAATATAGATTTAATCTGAGTGAAGTACTGAGC 2342
Db 2288 CTTCTGTGGTCTGAAACACTCAACAGAAATATAGATTTAATCTGAGTGAAGTACTGAGC 2347
QY 2343 TTTGCTCCACTAAACAGCTGAGCTTGGTCCACTAAACCAAGATGAAAATACAGAG 2402
Db 2348 TTTGCTCCACTAAACAGCTGAGCTTGGTCCACTAAACCAAGATGAAAATACAGAG 2407
QY 2403 TGACTCTATAACTCTGGTCTTTTAAAGAAAGCTGCCTTTTTCATTTTATAGACAAAATCTTTTC 2462
Db 2408 TGACTCTATAACTCTGGTCTTTTAAAGAAAGCTGCCTTTTTCATTTTATAGACAAAATCTTTTC 2467
QY 2463 AACGCTGAAATGTACCTAAATCTGGTCTTCTACCATATATGATGAGCTTCCCGAGGA 2522
Db 2468 AACGCTGAAATGTACCTAAATCTGGTCTTCTACCATATATGATGAGCTTCCCGAGGA 2527
QY 2523 TGAATGCTGTGTTTAAATTTTCAATAAGATAATTTGTCTACTAGCATTTTCAAGTGAATAG 2582
Db 2528 TGAATGCTGTGTTTAAATTTTCAATAAGATAATTTGTCTACTAGCATTTTCAAGTGAATAG 2587
QY 2583 TCTTCACTTTTAAATTTTATCTCTCTCTATAATAATGACATCCAGTTCATGGAGGA 2642

```
Db 2588 TCTTCACTTTTAAATATTATTCATCTTCTCTATAATAATGACATCCAGTTTCATGGAGCA 2647
Qy 2643 AAAAACAAGTTTCTTTGTTATCCTGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702
Db 2648 AAAAACAAGTTTCTTTGTTATCCTGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2707
Qy 2703 ACAGCATGAGGCTGTGAAGCTGACTGAGAAATCTCTGCTGAAGCCCTGTTCTGT 2762
Db 2708 ACAGCATGAGGCTGTGAAGCTGACTGAGAAATCTCTGCTGAAGCCCTGTTCTGT 2767
Qy 2763 TCTGCTCTCAACATGTATAATTTTATTTTGAATAATACATAATCTTTTCACTATG 2814
Db 2768 TCTGCTCTCAACATGTATAATTTTATTTTGAATAATACATAATCTTTTCACTATG 2819

RESULT 7
ID ADF82011 standard; DNA; 4450 BP.
XX ADF82011;
AC 26-FEB-2004 (first entry)
DT Leukaemia-related DNA sequence #2567.
DE Leukaemia-related DNA sequence #2567.
XX Cytostatic; Gene therapy; leukaemia; ss.
KW Unidentified.
XX WO2003039443-A2.
PN 15-MAY-2003.
XX 04-NOV-2002; 2002WO-EP012303.
XX 05-NOV-2001; 2001EP-00126244.
PR 30-APR-2002; 2002EP-00009759.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA (UYLU-) UNIV LUDWIG MAXIMILIANS.
PA (HAPE/) HAFERLACH T.
PA (SCHO/) SCHOCH C.
PA (KERN/) KERN W.
XX Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;
PI Eils R, Brors B, Mergenthaler S;
XX WPI; 2003-505037/47.
DR
XX Determining the subtype of leukemia cells and whether a patient sample
PT contains leukemia cells or other cells, useful for treating leukemia,
PT comprises determining the expression profile of a group of markers in a
PT patient sample.
XX Disclosure; SEQ ID NO 2567; 2938pp; English.
PS
XX The present invention relates to a method (M1) for determining the
CC subtype of leukaemia cells and whether a patient sample contains
CC leukaemia cells. The method comprises determining the expression profile
CC of a group of markers in a patient sample. The method is useful for
CC determining the presence of leukaemia cells, its types or subtypes, and
CC for the preparation of a medicament for treating leukaemia.
XX
SQ Sequence 4450 BP; 1307 A; 930 C; 888 G; 1305 T; 0 U; 20 Other;

Query Match 99.0%; Score 2802; DB 10; Length 4450;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2805; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CACGAGCGGGAGTTGGAGCGATACGATTTGTGTGTGAGAGGCGCAACGTGCGGATTC 62
Db 36 CAGTGGCGGAGTTGGAGCGATACGATTTGTGTGTGAGAGGCGCAACGTGCGGATTC 95
```

```
Qy 63 TGTGTAACCTGGAGGCAATTTCTAGCACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 122
Db 96 TGTGTAACCTGGAGGCAATTTCTAGCACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 155
Qy 123 GATGCTCTTCAATTCGGTGTCTCGGCCAGCCCGCAGCTTTGGCGTCTCGAGAAATGGATGGTC 182
Db 156 GATGCTCTTCAATTCGGTGTCTCGGCCAGCCCGCAGCTTTGGCGTCTCGAGAAATGGATGGTC 215
Qy 183 TTCAATAACCTCTTCAATTCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 242
Db 216 TTCAATAACCTCTTCAATTCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 275
Qy 243 CACTTCTTATGGAGAAACAGAGTCCCGAGTTCCTCTTTTGGATGTACCTTCTCTCTGC 302
Db 276 CACTTCTTATGGAGAAACAGAGTCCCGAGTTCCTCTTTTGGATGTACCTTCTCTCTGC 335
Qy 303 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAAGGCTTTTTCGATTTGATAA 362
Db 336 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAAGGCTTTTTCGATTTGATAA 395
Qy 363 CACAGAATCAAAAAGTTTCAGAAAGAGTCTTCAAAAGATGGATGGCTCACTGGAAATGC 422
Db 396 CACAGAATCAAAAAGTTTCAGAAAGAGTCTTCAAAAGATGGATGGCTCACTGGAAATGC 455
Qy 423 CGTCTTTGACCTGGCGCTGGGTTCTGTGGAACCTTAAACTTTGTACAGCAGCAGTGTATCA 482
Db 456 CGTCTTTGACCTGGCGCTGGGTTCTGTGGAACCTTAAACTTTGTACAGCAGCAGTGTATCA 515
Qy 483 AACGCCAAAATTTGGGACCGTAAAGCTGGTGAAGCTGATTTGGAACATGCAAAAGTCATCA 542
Db 516 AACGCCAAAATTTGGGACCGTAAAGCTGGTGAAGCTGATTTGGAACATGCAAAAGTCATCA 575
Qy 543 ATGCAAGCTCAAGTCAAGTTCCTTTCTAAGTTTGAAGAAAGCTGTATTTCTGTACGGGTGG 602
Db 576 ATGCAAGCTCAAGTCAAGTTCCTTTCTAAGTTTGAAGAAAGCTGTATTTCTGTACGGGTGG 635
Qy 603 AAGAGTGGCAACATATATGCTGGGATACCGAGTGCACAAAGAAAGATGGGTTTATAG 662
Db 636 AAGAGTGGCAACATATATGCTGGGATACCGAGTGCACAAAGAAAGATGGGTTTATAG 695
Qy 663 GCAAGTGAATCAAAATCAGTGGAGCTCAATACTCTCAGACAAGCAAAACCCCTTCAAAACC 722
Db 696 GCAAGTGAATCAAAATCAGTGGAGCTCAATACTCTCAGACAAGCAAAACCCCTTCAAAACC 755
Qy 723 CAGAGAAGAAACAGAAATTCAGAAAGCTGCTGCTCTCTCTGTGGAATTCAGAAAGTGTATC 782
Db 756 CAGAGAAGAAACAGAAATTCAGAAAGCTGCTGCTCTCTCTGTGGAATTCAGAAAGTGTATC 815
Qy 783 TGTGGTCTCTTCAAGACGAGAAATCCTTTAGTCTCAGCAGGAGCTGTGATGGGATAAT 842
Db 816 TGTGGTCTCTTCAAGACGAGAAATCCTTTAGTCTCAGCAGGAGCTGTGATGGGATAAT 875
Qy 843 CAAAGTATGGATTTTACGTAAGAAATTAATCTGCTTTATCGACAAGAACCCATAGCATCAA 902
Db 876 CAAAGTATGGATTTTACGTAAGAAATTAATCTGCTTTATCGACAAGAACCCATAGCATCAA 935
Qy 903 GTCTTTCTGTATCCAGGTAGCAGACTCGAAATCTGGATATTTCAAGTCTGATTTTGA 962
Db 936 GTCTTTCTGTATCCAGGTAGCAGACTCGAAATCTGGATATTTCAAGTCTGATTTTGA 995
Qy 963 TTCCACTGGCTCTACTTTTATTTGCTAATTCACAGCAGTAACATCTACATGTTTAAATAT 1022
Db 996 TTCCACTGGCTCTACTTTTATTTGCTAATTCACAGCAGTAACATCTACATGTTTAAATAT 1055
Qy 1023 GACTGGGTTGAAGAACTTCTCCAGTGGCTATTTTCAATGGACACCAGAACTCTACCTTTTA 1082
Db 1056 GACTGGGTTGAAGAACTTCTCCAGTGGCTATTTTCAATGGACACCAGAACTCTACCTTTTA 1115
Qy 1083 TGTAAATTCAGGCTTTAGTCCAGATGACCAAGTCTTTTGTAGTCAAGTGGCTCAAGTGAAGC 1142
Db 1116 TGTAAATTCAGGCTTTAGTCCAGATGACCAAGTCTTTTGTAGTCAAGTGGCTCAAGTGAAGC 1175
Qy 1143 TGCCTACATATGGAAGGTCTCCA.CACCTTGGCAACCTCTTCTACTGTGCTCTCTGGGTCTATTC 1202
```

1176	Db	 TGCCCTACATATGGAAGGCTCTCCACACCGCTGGCAACCTCTCTACTGTGCTCGGTGCAATTC	1235
1203	Qy	 TCAAGAGGTCACGTCGTCTGTGGTCTGCATCTTGACTTCACAAGAAGTTCGTCTACCTGTCTC	1262
1236	Db	 TCAAGAGGTCACGTCGTCTGTGGTCTGCATCTTGACTTCACAAGAAGTTCGTCTACCTGTCTC	1295
1263	Qy	 TGATGACAATACACATAAAAAATCTGGCGCTTTGAATAGAGGCTTATAGAGGAGAAAAACGAGAGG	1322
1296	Db	 TGATGACAATACACTAAAAATCTGGCGCTTTGAATAGAGGCTTATAGAGGAGAAAAACGAGAGG	1355
1323	Qy	 TGATAAACTTTCCACGCTGGGTTGGGCGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCCCT	1382
1356	Db	 TGATAAACTTTTCCACGCTGGGTTGGGCGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCCCT	1415
1383	Qy	 AGTAACAGTAGTAAACAGTAGTACCTCTCTGCCAAAAGCCGCCAGGCTTAAGTGCATATCC	1442
1416	Db	 AGTAACAGTAGTAACGAGTAGCAGAGTACTCTCTGCCAAAAGCCGCCAGGCTTAAGTGCATATCC	1475
1443	Qy	 ATCCAAATCTTTCCCGCTCATCCGCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCT	1502
1476	Db	 ATCCAAATCTTTCCCGCTCATCCGCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCT	1535
1503	Qy	 TCCTTCAAAATCTCCTACGTTCTCTATTAAAACTCTCTTGCCCAAGGCCCGGTCTCCCAT	1562
1536	Db	 TCCTTCAAAATCTCCTACGTTCTCTATTAAAACTCTCTTGCCCAAGGCCCGGTCTCCCAT	1595
1563	Qy	 CAACAGAAAGGCTCTGTCCTCCGTCCTCTCCCAAGCCACTTCATCTTTTCAAGATGTC	1622
1596	Db	 CAACAGAAAGGCTCTGTCCTCCGTCCTCTCCCAAGCCACTTCATCTTTTCAAGATGTC	1655
1623	Qy	 GATTAGAAACTGGGTGACCCGGAACACCTTCTCATCAACACCCATCACTCCACTGTCTTC	1682
1656	Db	 GATTAGAAACTGGGTGACCCGGAACACCTTCTCATCAACACCCATCACTCCACTGTCTTC	1715
1683	Qy	 GGAGACAAGATCATGTCCTCGAGAAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCCA	1742
1716	Db	 GGAGACAAGATCATGTCCTCGAGAAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCCA	1775
1743	Qy	 AGCAGAGGCTTGCTCTCAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT	1802
1776	Db	 AGCAGAGGCTTGCTCTCAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT	1835
1803	Qy	 GGAGAGTGTGAAAACAAAAGTGTGTGAAGAGTGTGTAACTGTGTGACTCAGCTTGATGGCCA	1862
1836	Db	 GGAGAGTGTGAAAACAAAAGTGTGTGAAGAGTGTGTAACTGTGTGACTCAGCTTGATGGCCA	1895
1863	Qy	 AGTTGAAAAATCTTCAATTTGGATCTGTCTGCCTTGTGTGTAACAGGAAGACCTTAGTAA	1922
1896	Db	 AGTTGAAAAATCTTCAATTTGGATCTGTCTGCCTTGTGTGTAACAGGAAGACCTTAGTAA	1955
1923	Qy	 GGACTCTCTAGGCTCTACCAAAATCAAGCAAAATTTGAAGAGGCTGGTACCAAGTATCTCAGA	1982
1956	Db	 GGACTCTCTAGGCTCTACCAAAATCAAGCAAAATTTGAAGAGGCTGGTACCAAGTATCTCAGA	2015
1983	Qy	 GCCTCCGCTCTCTATCAGTCCGATGTGTCTCAGAAAAGCTGTGTGAACGCTACTCTTCTCTTT	2042
2016	Db	 GCCTCCGCTCTCTATCAGTCCGATGTGTCTCAGAAAAGCTGTGTGAACGCTACTCTTCTCTTT	2075
2043	Qy	 GAGACCTTTGTGGAGAGGGCTGTAATGGTAGGCAAAAGAGAAATAGTTTCCCAGAGAAATAA	2102
2076	Db	 GAGACCTTTGTGGAGAGGGCTGTAATGGTAGGCAAAAGAGAAATAGTTTCCCAGAGAAATAA	2135
2103	Qy	 AAACTGGTTGTTGGCCATGGCGCAAAACGGAGGCTGAGAAATCCATCTCCACGAAGTCC	2162
2136	Db	 AAACTGGTTGTTGGCCATGGCGCAAAACGGAGGCTGAGAAATCCATCTCCACGAAGTCC	2195
2163	Qy	 GTCAATCCAGACACCCAAATTCAGAGACAGAGCGGHAAGACATTTGCCAAGCCCGGTTCAC	2222
2196	Db	 GTCAATCCAGACACCCAAATTCAGAGACAGAGCGGHAAGACATTTGCCAAGCCCGGTTCAC	2255
2223	Qy	 CATCAGCCCGAGCTCCATGATGAGGAAAAATCTGCACTACTTCCATAGAAAGTCCCAGGAGGA	2282

Db	2256	CATCACGCCCAGCTCCATGATGAGGAAAATCTGCACATACTCTTCATAGAAAATGCCCAGAGGA	2311
Qy	2283	CTTCTGTGGTCCTGAAACACCTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTTACTGAGC	2342
Db	2316	CTTCTGTGGTCCTGAAACACCTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTTACTGAGC	2375
Qy	2343	TTTGGTCCACTAAAAACAAGCTGAGCTTTGGTGCACCTAAAAACAAGATGAAAAATACAAGAG	2402
Db	2376	TTTGGTCCACTAAAAACAAGCTGAGCTTTGGTGCACCTAAAAACAAGATGAAAAATACAAGAG	2435
Qy	2403	TGACTCTATAACTCTGGTCTTTTAAGAAAGCTGGCTTTTTCATATTTTAGACAAAAATCTTTTC	2462
Db	2436	TGACTCTATAACTCTGGTCTTTTAAGAAAGCTGGCTTTTTCATATTTTAGACAAAAATCTTTTC	2495
Qy	2463	AACGCTGAAATGTACCTAAATCTGGTCTTCTACTACCATAAATGATATGACAGCTTCCCGAGGA	2522
Db	2496	AACGCTGAAATGTACCTAAATCTGGTCTTCTACTACCATAAATGATATGACAGCTTCCCGAGGA	2555
Qy	2523	TGAATGCTGTGTTTAAATTTCAFAAAGTAAATTTGTGCACCTCTAGCAATTTTGAATGAATAG	2582
Db	2556	TGAATGCTGTGTTTAAATTTCAFAAAGTAAATTTGTGCACCTCTAGCAATTTTGAATGAATAG	2615
Qy	2583	TCCTTCACCTTTTAAATTTATTCATCTTCTCTATAATATGACATCCAGTTTCATGGAGGCA	2642
Db	2616	TCCTTCACCTTTTAAATTTATTCATCTTCTCTATAATATGACATCCAGTTTCATGGAGGCA	2675
Qy	2643	AAAAACAAGTTTCTTGTTATCCTGAAACTTTCTATGCTCAGTGGAAAAGTATCTGCCAGCC	2702
Db	2676	AAAAACAAGTTTCTTGTTATCCTGAAACTTTCTATGCTCAGTGGAAAAGTATCTGCCAGCC	2735
Qy	2703	ACAGCATGAGGCTGTGAAGGCTGACCTGAGAAATCTCTGCTGAAAGACCCCTGGTTCTGT	2762
Db	2736	ACAGCATGAGGCTGTGAAGGCTGACCTGAGAAATCTCTGCTGAAAGACCCCTGGTTCTGT	2795
Qy	2763	TCTGCCTCCACATGTATATATTTTATTTGAAATACATAATCTTTTCACTA	2812
Db	2796	TCTGCCTCCACATGTATATATTTTATTTGAAATACATAATCTTTTCACTA	2845
RESULT 8			
ID	AAD63212 standard; DNA; 4221 BP.		
XX	AAD63212;		
XX	12-FEB-2004 (first entry)		
XX	Human DNA #37 used in the method for diagnosing cancer.		
DE	Human; breast cancer; metastasis; differential modulation; therapy; ds.		
KW	Homo sapiens.		
OS	US2003190656-A1.		
XX	09-OCT-2003.		
XX	21-MAR-2003; 2003US-00393590.		
XX	29-MAR-2002; 2002US-0368789P.		
XX	(WANG/) WANG Y.		
XX	Wang Y;		
XX	WPI; 2003-831621/77.		
XX	Prognosticating metastasis in a breast cancer patient comprises		
PT	identifying differential modulation of each gene relative to the		
PT	expression of the same genes in a normal population in combination of		
PT	genes.		
XX	Claim 1; Page 122-124; 0pp; English.		

XX The present invention relates to a method of prognosticating metastasis
CC in a breast cancer patient involves identifying differential modulation
CC of each gene relative to the expression of the same genes in a normal
CC population in combination of genes. The invention is useful for
CC prognosticating breast cancer in a patient. The present sequence is human
CC DNA used in the method for diagnosing cancer
XX

SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;

Query Match 98.5%; Score 2789.8; DB 10; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 CGATTAACGATTGTGTTGTGAGAGCGCAAGCTGCGATTTCTGCTGAACTTGGAGGCATT 81
DB |||||
1 CGATTAACGATTGTGTTGTGAGAGCGCAAGCTGCGATTTCTGCTGAACTTGGAGGCATT 60
QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCGACCCCTGATGCTCTTCAATTCGGTG 141
DB |||||
61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCGACCCCTGATGCTCTTCAATTCGGTG 120
QY 142 CTCGCCACGCCAGCTTGGCGTCTCGAGAAATGATGCTTTCACAATACCCCTCTTCAA 201
DB |||||
121 CTCGCCACGCCAGCTTGGCGTCTCGAGAAATGATGCTTTCACAATACCCCTCTTCAA 180
QY 202 TCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA 261
DB |||||
181 TCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA 240
QY 262 GGAGTCCCAAGTTCCTCTTTGGATGTACCTTCTCTGCTCCCAATATGGACATGTA 321
DB |||||
241 GGAGTCCCAAGTTCCTCTTTGGATGTACCTTCTCTGCTCCCAATATGGACATGTA 300
QY 322 CTAGCAGTTGCCAATGAAGAGGCTTTGTTGATTGTATTAACACAGATCAACAAGTTTC 381
DB |||||
301 CTAGCAGTTGCCAATGAAGAGGCTTTGTTGATTGTATTAACACAGATCAACAAGTTTC 360
QY 382 AGAAAGAGTGTCTCAAGAATGSGATGGCTCACTGGAATGCGCTTTTGAAGTGGCCTGG 441
DB |||||
361 AGAAAGAGTGTCTCAAGAATGSGATGGCTCACTGGAATGCGCTTTTGAAGTGGCCTGG 420
QY 442 GTTCTGTGTAACTTAACTGTTACAGCAGCAGTGTATCAACAGCCCAATTTTGGAC 501
DB |||||
421 GTTCTGTGTAACTTAACTGTTACAGCAGCAGTGTATCAACAGCCCAATTTTGGAC 480
QY 502 GTAAAGCTGGTGAAGTGAATGGAACATGCAAGGTCAATGAGCCTCAAGTCAGTT 561
DB |||||
481 GTAAAGCTGGTGAAGTGAATGGAACATGCAAGGTCAATGAGCCTCAAGTCAGTT 540
QY 562 GCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGACGGGTGGAAGAGATGGCAACATTATG 621
DB |||||
541 GCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGACGGGTGGAAGAGATGGCAACATTATG 600
QY 622 GTCTGGGTACAGGTGCACAAAAGATGGTTTTATAGCAGTGAATCAATCAGT 681
DB |||||
601 GTCTGGGTACAGGTGCACAAAAGATGGTTTTATAGCAGTGAATCAATCAGT 660
QY 682 GGAGCTCAAAATACCTCAGCAAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATTC 741
DB |||||
661 GGAGCTCAAAATACCTCAGCAAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATTC 720
QY 742 AAAGGACTTGTCTTCTGTGATTTCAGCAAAAGTGTACTGTGGTCTCTTTTCAAGAC 801
DB |||||
721 AAAGGACTTGTCTTCTGTGATTTCAGCAAAAGTGTACTGTGGTCTCTTTTCAAGAC 780
QY 802 GAGATACCTTACTCTCAGCAGGAGTGGATGGATTAATCAAGATATGGATTATCGT 861
DB |||||
781 GAGATACCTTACTCTCAGCAGGAGTGGATGGATTAATCAAGATATGGATTATCGT 840
QY 862 AAGAATTATCTGCTTATCGAAGAACCCATAGCATCCAAGTCTTCTGCTGACCCAGT 921
DB |||||
841 AAGAATTATCTGCTTATCGAAGAACCCATAGCATCCAAGTCTTCTGCTGACCCAGT 900

QY 922 AGAGCACTCGAANAACCTGGATATTCAAGTCTGATTTTGGATTCCACTGGCTCTACTTTA 981
DB |||||
901 AGAGCACTCGAANAACCTGGATATTCAAGTCTGATTTTGGATTCCACTGGCTCTACTTTA 960
QY 982 TTTGCTAATTGACAGACGATTAACATCTACATGTTTAAATATGACTGGTGTGAAGACTTCT 1041
DB |||||
961 TTTGCTAATTGACAGACGATTAACATCTACATGTTTAAATATGACTGGTGTGAAGACTTCT 1020
QY 1042 CCAGTGGCTATTCTTCAATGGACACCAGAACTCTACCTTTTATGTPAAATCCAGCTTAGT 1101
DB |||||
1021 CCAGTGGCTATTCTTCAATGGACACCAGAACTCTACCTTTTATGTPAAATCCAGCTTAGT 1080
QY 1102 CCAGATGACACAGTTTTTATGTCAGTGGCTCAAGTGAATGAAGCTGCTACATATGAAGGTC 1161
DB |||||
1081 CCAGATGACACAGTTTTTATGTCAGTGGCTCAAGTGAATGAAGCTGCTACATATGAAGGTC 1140
QY 1162 TCCACACCTGGCAACCTCTCTACTCTCTCTGGGTCAATCTCAAGAGGTCACTCTGTGTG 1221
DB |||||
1141 TCCACACCTGGCAACCTCTCTACTCTCTCTGGGTCAATCTCAAGAGGTCACTCTGTGTG 1200
QY 1222 TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTTCTGATGAACAATACACTAAAA 1281
DB |||||
1201 TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTTCTGATGAACAATACACTAAAA 1260
QY 1282 ATCTGGCGCTTGATAGAGCTTAGAGGAAACACAGAGGTGATAACTTTCCACCGGTG 1341
DB |||||
1261 ATCTGGCGCTTGATAGAGCTTAGAGGAAACACAGAGGTGATAACTTTCCACCGGTG 1320
QY 1342 GGTTCGGCTCTCTCAGAAAGAAAAGAGTCAAGACCTGGCTTAGTAAACAAGTAAACGAGTAGC 1401
DB |||||
1321 GGTTCGGCTCTCTCAGAAAGAAAAGAGTCAAGACCTGGCTTAGTAAACAAGTAAACGAGTAGC 1380
QY 1402 CAGAGTACTCTCTGCAAAAGCCCCCAGAGTAAAGTGCAATCCATCAATTTCTTCCCGTCA 1461
DB |||||
1381 CAGAGTACTCTCTGCAAAAGCCCCCAGAGTAAAGTGCAATCCATCAATTTCTTCCCGTCA 1440
QY 1462 TCCGAGTGTGTGCCCCAGCTGTCTGGAGACCTCTCTCTCTCTTCAATATCTCTTACG 1521
DB |||||
1441 TCCGAGTGTGTGCCCCAGCTGTCTGGAGACCTCTCTCTCTCTTCAATATCTCTTACG 1500
QY 1522 TTCTCTATTTAAACCTCTCTGCAAGCCCGGTCTCCCATCAACAGAGAGGTCTGTCTG 1581
DB |||||
1501 TTCTCTATTTAAACCTCTCTGCAAGCCCGGTCTCCCATCAACAGAGAGGTCTGTCTG 1560
QY 1582 TCCTCCGTCTCTCCCAAGCCACCTTCTATCTTTTCAAGATGTCGATTAGAAACTGGGTGACC 1641
DB |||||
1561 TCCTCCGTCTCTCCCAAGCCACCTTCTATCTTTTCAAGATGTCGATTAGAAACTGGGTGACC 1620
QY 1642 CGAACCTTCTCTCATCACCCCATCATCTCCACCTGCTTCGGAGACCAAGATCATGTCT 1701
DB |||||
1621 CGAACCTTCTCTCATCACCCCATCATCTCCACCTGCTTCGGAGACCAAGATCATGTCT 1680
QY 1702 CCGAGAAAAGCCCTTATTTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1761
DB |||||
1681 CCGAGAAAAGCCCTTATTTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1740
QY 1762 TCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1821
DB |||||
1741 TCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1800
QY 1822 TGTGTGAAGAGTGTGTAAGTGTGACTGAGCTTGAATGAGTGAATTTGAAATCTTCATTG 1881
DB |||||
1801 TGTGTGAAGAGTGTGTAAGTGTGACTGAGCTTGAATGAGTGAATTTGAAATCTTCATTG 1860
QY 1882 GATCTGTCTCCCTTGTGTTAAACAGAAACCTTAGTAGAGGACTCTCTAGAGTCTTAC 1941
DB |||||
1861 GATCTGTCTCCCTTGTGTTAAACAGAAACCTTAGTAGAGGACTCTCTAGAGTCTTAC 1920
QY 1942 AAATCAAGCAAAATTAAGAGGAGTGGTACAGTATCTCAGAGCCTTCGGTCTCTTATCAGT 2001
DB |||||
1921 AAATCAAGCAAAATTAAGAGGAGTGGTACAGTATCTCAGAGCCTTCGGTCTCTTATCAGT 1980


```
QY 2002 CCGTATGCTTCAGAAAGCTGTGGAAGCTACCTCTCTCTTTCAGACCTTGTGTGAGAGGG 2061
Db 1981 CCGTATGCTTCAGAAAGCTGTGGAAGCTACCTCTCTCTTTCAGACCTTGTGTGAGAGGG 2040
QY 2062 TCTGAAATGGTATAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAAAAGTGTGTGTGCGCATG 2121
Db 2041 TCTGAAATGGTATAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAAAAGTGTGTGTGCGCATG 2100
QY 2122 CGAGCCAAACGAGGCTGAGAAATCCATCTCCACGAGGCTCCGTCATCCGAGACCCCAAT 2181
Db 2101 CGAGCCAAACGAGGCTGAGAAATCCATCTCCACGAGGCTCCGTCATCCGAGACCCCAAT 2160
QY 2182 TCCAGAGACAGAGCGGAAAGCAATTCGCAAGCCGCTCACCATCAGCGCCAGCTCCCATG 2241
Db 2161 TCCAGAGACAGAGCGGAAAGCAATTCGCAAGCCGCTCACCATCAGCGCCAGCTCCCATG 2220
QY 2242 AGGAAATCTGCACATACCTTCCATAGAAAGTCCGAGGAGGACTTCTGTGGTCTCTGAAACAC 2301
Db 2221 AGGAAATCTGCACATACCTTCCATAGAAAGTCCGAGGAGGACTTCTGTGGTCTCTGAAACAC 2280
QY 2302 TCAACAGAAATTAAGATCTTAATCTGAGTGAGTACTGAGCTTTGGTCCACTAAACACAG 2361
Db 2281 TCAACAGAAATTAAGATCTTAATCTGAGTGAGTACTGAGCTTTGGTCCACTAAACACAG 2340
QY 2362 CTGAGCTTTGGTCCACTAAACAGATGAAATACAGAGTGACTCTATAACTCTGGTC 2421
Db 2341 CTGAGCTTTGGTCCACTAAACAGATGAAATACAGAGTGACTCTATAACTCTGGTC 2400
QY 2422 TTTAAGAAAGCTGCCCTTTTCATTTTTAGACAAAATCTTTTCAACGCTGAAATGTACCTAA 2481
Db 2401 TTTAAGAAAGCTGCCCTTTTCATTTTTAGACAAAATCTTTTCAACGCTGAAATGTACCTAA 2460
QY 2482 TCTGGTCTACTACCAATATGATATGAGCTTCGAGCTTCGAGGATGAATGCTGTGTAAAT 2541
Db 2461 TCTGGTCTACTACCAATATGATATGAGCTTCGAGGATGAATGCTGTGTAAAT 2520
QY 2542 TCATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATAT 2601
Db 2521 TCATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATAT 2580
QY 2602 TCATCTCTCTATAATAATGACATCCAGTTCATGAGGCAAAAAACAAGTTCTTGTGTA 2661
Db 2581 TCATCTCTCTATAATAATGACATCCAGTTCATGAGGCAAAAAACAAGTTCTTGTGTA 2640
QY 2662 TCCGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGCCCTGTGAA 2721
Db 2641 TCCGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGCCCTGTGAA 2700
QY 2722 GGCTGACTGAGAAATCTCTGCTGAAGACCCCTGTTCTGTTCTGCTCCACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCTCTGCTGAAGACCCCTGTTCTGTTCTGCTCCACATGTATA 2760
QY 2782 ATTTTATTGAAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTGAAATACATAATCTTTTCACTATG 2793
```

RESULT 9

AAD62813

ID AAD62813 standard; DNA; 4221 BP.

AC AAD62813;

XX AAD62813;

XX 12-FEB-2004 (first entry)

DE Human DNA #26 used in the method for diagnosing cancer.

XX Human; cancer; differential modulation; gene expression profile; ds.

XX Homo sapiens.

XX OS

XX PN

XX US2003194733-A1.

XX

```
PD 16-OCT-2003.
XX
PF 21-MAR-2003; 2003US-00393567.
XX
PR 29-MAR-2002; 2002US-0368667P.
XX (WANG/) WANG Y.
XX
PI Wang Y;
XX
DR WPI; 2003-844450/78.
XX
PT Diagnosing cancer comprises identifying differential modulation of each
PT gene (relative to the expression of the same genes in a normal
PT population) in a combination of genes.
XX
PS Example 3; Page 122-124; Opp; English.
XX
CC The invention relates to a method for diagnosing cancer. The method
CC comprising identifying differential modulation of each gene (relative to
CC the expression of the same genes in a normal population) in a combination
CC of genes. The method, diagnostic portfolio and the kit are useful in
CC diagnosing cancer. The method may also be used for determining gene
CC expression profiles. The present sequence is human DNA used to illustrate
CC the method of the invention
XX
SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;

Query Match 98.5%; Score 2789.8; DB 10; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 CGATAACGATTTGTGTGTGTGAGAGCGCAACGTCGGATTTCTGCTGAACCTTGGAGGCATT 81
Db 1 CGATAACGATTTGTGTGTGTGAGAGCGCAACGTCGGATTTCTGCTGAACCTTGGAGGCATT 60
QY 82 TCTAGACATTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 141
Db 61 TCTAGACATTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 120
QY 142 CTCGCCACGCCCCAGCTTGGGCTCTGAGAAATGGAATGGTCTTTCACAAATACCCCTTTCAA 201
Db 121 CTCGCCACGCCCCAGCTTGGGCTCTGAGAAATGGAATGGTCTTTCACAAATACCCCTTTCAA 180
QY 202 TCCCTTCTGACTGGTATCAGTGAGTGATGATGACACACTTCTTATGGAGAAACA 261
Db 181 TCCCTTCTGACTGGTATCAGTGAGTGATGATGACACACTTCTTATGGAGAAACA 240
QY 262 GGAGTCCAGTTCCTCCCTTTTGGATGTACCTTCTCTCTGCTCCCAATATGGAACATGTA 321
Db 241 GGAGTCCAGTTCCTCCCTTTTGGATGTACCTTCTCTCTGCTCCCAATATGGAACATGTA 300
QY 322 CTAGCAGTTGCCAATGAAAGAGGCTTTGTTCCGATTTGATATAACACAGAAATCAAAAGTTTC 381
Db 301 CTAGCAGTTGCCAATGAAAGAGGCTTTGTTCCGATTTGATATAACACAGAAATCAAAAGTTTC 360
QY 382 AGAAGAAGTCTTCAAGAATGATGCTCAGTGGAATGCGTCTTTCAGCTGCGCTGG 441
Db 361 AGAAGAAGTCTTCAAGAATGATGCTCAGTGGAATGCGTCTTTCAGCTGCGCTGG 420
QY 442 GTTCTGCTGAACTTAAACTTTTACAGCAGCAGTGATCAAAACAGCCAAATTTTGGGAC 501
Db 421 GTTCTGCTGAACTTAAACTTTTACAGCAGCAGTGATCAAAACAGCCAAATTTTGGGAC 480
QY 502 GTAAAAAGCTGCTGAGCTGATTTGGAAATGCAATGCAAGGCTCAATGAGGCTCAAGTCAGTT 561
Db 481 GTAAAAAGCTGCTGAGCTGATTTGGAAATGCAATGCAAGGCTCAATGAGGCTCAAGTCAGTT 540
QY 562 GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGATGGCAACATTATG 621
Db 541 GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGATGGCAACATTATG 600
QY 622 GTCTGGGATACCAGGTGCAACAAAAAAGATGGGTTTTTATAGGCAAGTGAATCAAAATCAGT 681
```

|||||
601 GTCTGGGATACAGGTGCACAAAAAAGATGGTTTTTATAGGCAAGTGAATCAATCAGT 660
QY
682 GGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCAAAACCCAAAGAGAAAACAGAAATTC 741
Db
661 GGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCAAAACCCAAAGAGAAAACAGAAATTC 720
QY
742 AAAGGACCTTGTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTCAAGAC 801
Db
721 AAAGGACCTTGTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTCAAGAC 780
QY
802 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGATGGGATAATCAAAAGTATGGGATTAAGT 861
Db
781 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGATGGGATAATCAAAAGTATGGGATTAAGT 840
QY
862 AAGAAATTACTGCTTTATCGACAAGACCCATAGCATCCAAGTCTTTCTCTGTACCCAGGT 921
Db
841 AAGAAATTACTGCTTTATCGACAAGACCCATAGCATCCAAGTCTTTCTCTGTACCCAGGT 900
QY
922 AGCAGCACTCGAAACTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
Db
901 AGCAGCACTCGAAACTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
QY
982 TTTTGCTAATTGCACAGACGATAACATCTACATGTTTAAATATGACTGGGTTCGAAGACTTCT 1041
Db
961 TTTTGCTAATTGCACAGACGATAACATCTACATGTTTAAATATGACTGGGTTCGAAGACTTCT 1020
QY
1042 CCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1101
Db
1021 CCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1080
QY
1102 CCAGATGACCAAGTTTTATGTCAGTGGCTCAAGTGAAGCTGCCTACATATGGAGGTC 1161
Db
1081 CCAGATGACCAAGTTTTATGTCAGTGGCTCAAGTGAAGCTGCCTACATATGGAGGTC 1140
QY
1162 TCCACACCTCGCAACCTCTACTGCTCTCTGGGTCAATCTCAAGAGGTCACGCTCTGTG 1221
Db
1141 TCCACACCTCGCAACCTCTACTGCTCTCTGGGTCAATCTCAAGAGGTCACGCTCTGTG 1200
QY
1222 TGCTGGTGTCCATCTGACTTCAAAAGATTTGCTACCTGTTCTGTATGACAATAACATAAAA 1281
Db
1201 TGCTGGTGTCCATCTGACTTTCACAAAGATTTGCTACCTGTTCTGTATGACAATAACATAAAA 1260
QY
1282 ATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACCCAGAGGATGATAAACTTTCCAGGTC 1341
Db
1261 ATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACCCAGAGGATGATAAACTTTCCAGGTC 1320
QY
1342 GGTTCGGGCTCTCAGAAGAAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAACGAGTAGC 1401
Db
1321 GGTTCGGGCTCTCAGAAGAAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAACGAGTAGC 1380
QY
1402 CAGAGTACTCTCTGCAAAAGCCCAAGGGTAAAGTGAATCCATCCAATTTCTTCCCGGTCA 1461
Db
1381 CAGAGTACTCTCTGCAAAAGCCCAAGGGTAAAGTGAATCCATCCAATTTCTTCCCGGTCA 1440
QY
1462 TCCGAGCTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATACCTCTAGC 1521
Db
1441 TCCGAGCTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATACCTCTAGC 1500
QY
1522 TTTCTCTATTAACCTCTCTGCAAGCCCGGCTCTCCCATCAACAGAGAGGCTCTGTC 1581
Db
1501 TTTCTCTATTAACCTCTCTGCAAGCCCGGCTCTCCCATCAACAGAGAGGCTCTGTC 1560
QY
1582 TCCTCCGCTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGATTTAGAACTGGGTGACC 1641
Db
1561 TCCTCCGCTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGATTTAGAACTGGGTGACC 1620
QY
1642 CGAACACCTTCTCTATCAACCCATCACTCCACTGCTTGGAGACCAAGATCAATGCT 1701
Db
1621 CGAACACCTTCTCTATCAACCCATCACTCCACTGCTTGGAGACCAAGATCAATGCT 1680
QY
1702 CCGAGAAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGCTCTGAG 1761
|||||

Db
1681 CCGAGAAAAGCCCTTATTTCTGTGAGCCAGAAGTCAATCCCAAGCAGAGGCTTGCTCTGAG 1740
QY
1762 TCTAGAAATAGAGTAAAGAGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAAG 1821
Db
1741 TCTAGAAATAGAGTAAAGAGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAAG 1800
QY
1822 TGTGTGAAGAGTTCTAACTGTGTGACTGTAGCTTGCATGCGCAAGTTGAAAATCTTTCATTG 1881
Db
1801 TGTGTGAAGAGTTGTAACTGTGTGACTGTAGCTTGTATGCGCCNAAGTTGAAAATCTTTCATTG 1860
QY
1882 GATCTGTGCTGCTTGTGCTTAACAGGAAGACCTTAGTAAGGACTCTCTAGGTCCTTACC 1941
Db
1861 GATCTGTGCTGCTTGTGCTTAACAGGAAGACCTTAGTAAGGACTCTCTAGGTCCTTACC 1920
QY
1942 AAATCAAGCAAAAATTTGAAGGAGCTGTACCGATATCTCAGAGCCTCGTCTCCTATCAGT 2001
Db
1921 AAATCAAGCAAAAATTTGAAGGAGCTGTACCGATATCTCAGAGCCTCGTCTCCTATCAGT 1980
QY
2002 CCGTATCTTTCAGAAAGCTGTGGAAACGCTACCTCTCTCTTTTGAGACCTTTGGAGAGAGG 2061
Db
1981 CCGTATCTTTCAGAAAGCTGTGGAAACGCTACCTCTCTCTTTTGAGACCTTTGGAGAGAGG 2040
QY
2062 TCTCAAAATGCTAGGCAAGAGAAATAGTTCCCGAGAGAAATAAAAACCTGGTTGTCGCCATG 2121
Db
2041 TCTCAAAATGCTAGGCAAGAGAAATAGTTCCCGAGAGAAATAAAAACCTGGTTGTCGCCATG 2100
QY
2122 GCAGCCAAAACGGAAGGCTGAGAAATCATCTCCAGGAAGTCCGTCATGCCAGACACCCAAAT 2181
Db
2101 GCAGCCAAAACGGAAGGCTGAGAAATCATCTCCAGGAAGTCCGTCATGCCAGACACCCAAAT 2160
QY
2182 TCCAGGAGACAGAGCGGAAGAGACAATGCCAAGCCCGGTCAACATCAGCCAGCTCCATG 2241
Db
2161 TCCAGGAGACAGAGCGGAAGAGACAATGCCAAGCCCGGTCAACATCAGCCAGCTCCATG 2220
QY
2242 AGGAAAATCTGCACTACTTCCATAGAAAAGTCCAGGAGGACTCTCTGTGGTCTCTGAACAC 2301
Db
2221 AGGAAAATCTGCACTACTTCCATAGAAAAGTCCAGGAGGACTCTCTGTGGTCTCTGAACAC 2280
QY
2302 TCAACAGAAATATAGATTTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTTAAACAG 2361
Db
2281 TCAACAGAAATATAGATTTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTTAAACAG 2340
QY
2362 CTGAGCTTTGTGCTCCACTTAAACAAAGATGAAAATGAAAATGAAAATGAAAATGAAAATG 2421
Db
2341 CTGAGCTTTGTGCTCCACTTAAACAAAGATGAAAATGAAAATGAAAATGAAAATGAAAATG 2400
QY
2422 TTTAAGAAAAGCTGCTTTTCACTTTTATAGACAAAATCTTTTCAACGCTGAAAATGACCTAA 2481
Db
2401 TTTAAGAAAAGCTGCTTTTCACTTTTATAGACAAAATCTTTTCAACGCTGAAAATGACCTAA 2460
QY
2482 TCTGGTCTCTACTACATAATGATATGAGCTTCCCGAGGATGAATGCTGTGTTTAAAT 2541
Db
2461 TCTGGTCTCTACTACATAATGATATGAGCTTCCCGAGGATGAATGCTGTGTTTAAAT 2520
QY
2542 TCATAAGTAAATTTGTCACTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2601
Db
2521 TCATAAGTAAATTTGTCACTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2580
QY
2602 TCATCTCTCTATATATATGACATCCAGATTTCCAGGAGGCAAAAACAAAGTTCTTGTGTA 2661
Db
2581 TCATCTCTCTATATATATGACATCCAGATTTCCAGGAGGCAAAAACAAAGTTCTTGTGTA 2640
QY
2662 TCCTGAAACTTTTCTATGCTAGTGGAAGATATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
Db
2641 TCCTGAAACTTTTCTATGCTAGTGGAAGATATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
QY
2722 GGCTGACTGGAATCTCTGCTGAAGACCCCTGCTTCTGTTCTGCTCCCAACATGTATA 2781
Db
2701 GGCTGACTGGAATCTCTGCTGAAGACCCCTGCTTCTGTTCTGCTCCCAACATGTATA 2760
QY
2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db
2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793
|||||

RESULT 10
ABX95498
ID ABX95498 standard; cDNA; 4221 BP.
AC ABX95498;
DT 01-JUL-2003 (first entry)
XX Human retinoic acid-regulated nuclear matrix-associated protein cDNA.
KW Nucleic acid identification; T cell receptor; T cell; tumour antigen;
KW antigen presenting cell; HLA; nucleic acid library; ELISPOT assay;
KW enzyme-linked immunospot; T cell receptor binding epitope;
KW cancer associated antigen; ramp; human; polytope; gene; ss;
KW retinoic acid-regulated nuclear matrix-associated protein.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 103..2295
FT /tag= a
FT /product= "ramp"
FT /note= "retinoic acid-regulated nuclear matrix-associated
protein"
PN US2003003485-A1.
XX
PD 02-JAN-2003.
XX
PF 14-MAY-2002; 2002US-00145396.
XX
PR 15-MAY-2001; 2001US-0291125P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Uenaka A, Nakayama E;
XX
DR WPI; 2003-416593/39.
DR P-PSDB; AB009611.
XX
PT Identifying nucleic acids encoding epitopes that bind T cell receptors on
PT T cells when presented by human leukocyte antigen by using an ELISPOT
PT assay that detects factors secreted by T cell in response to the receptor
PT binding.
XX
PS Disclosure; Page 33-35; 51pp; English.
XX
CC The invention describes a method of identifying nucleic acid encoding an
CC epitope that binds to a T cell receptor on T cell. The method involves
CC coculturing antigen presenting cells that contain a nucleic acid library
CC and express HLA molecules presenting epitopes, with a T cell having a T
CC cell receptor that binds the epitope so that the T cell receptor binds
CC the epitope encoded by the nucleic acid library. Detecting a factor
CC secreted by the T cell in response to T cell receptor binding, using
CC ELISPOT (enzyme-linked immunospot) assay, and correlating factor
CC secretion with presence of a nucleic acid encoding the epitope. The
CC method is useful for identifying a nucleic acid molecule encoding the
CC epitope that specifically binds to a T cell receptor on a T cell when
CC presented by an HLA molecule. A second method described in the invention
CC is useful for identifying antigens that specifically bind to a T cell
CC receptor on a T cell. The method is useful for identifying cancer
CC associated antigens. This sequence encodes the human retinoic acid-
CC regulated nuclear matrix-associated protein (ramp) peptides of which can
CC be combined with peptides of one or more other cancer-associated antigens
CC to form polytopes
XX
SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;
Query Match 98.5%; Score 2789.8; DB 10; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 CGATAACGATTGTGTTGTGAGAGCGCAACGTCGCGATTCTCTGCTGAACCTTGGAGGCATT 81
DB 1 CGATAACGATTGTGTTGTGAGAGCGCAACGTCGCGATTCTCTGCTGAACCTTGGAGGCATT 60
QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 141
DB 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 120
QY 142 CTCGCCAGCCCGACCTTGGCGTCTGAGAAATGGATGGTCTTCAATACCTCTTCA 201
DB 121 CTCGCCAGCCCGACCTTGGCGTCTGAGAAATGGATGGTCTTCAATACCTCTTCA 180
QY 202 TCCCTTCTGACTGGTATATCAGTGCAAGTGAATGATGAACACACTTCTTATGGAGAAACA 261
DB 181 TCCCTTCTGACTGGTATATCAGTGCAAGTGAATGATGAACACACTTCTTATGGAGAAACA 240
QY 262 GGAGTCCCAGTTCCTCTCTTTGGATGTACCTTCTCTCTGCCAATATGGAACATGTA 321
DB 241 GGAGTCCCAGTTCCTCTCTTTGGATGTACCTTCTCTCTGCCAATATGGAACATGTA 300
QY 322 CTAGCAGTTGCCAATGAAGAGGCTTTGTTGATGTATATACACAGAAATCAAAAGTTTC 381
DB 301 CTAGCAGTTGCCAATGAAGAGGCTTTGTTGATGTATATACACAGAAATCAAAAGTTTC 360
QY 382 AGAAGAAGTCTCTCAAGAATGATGGCTCACTGGAATGCGTCTTTGACCTGGCCTGG 441
DB 361 AGAAGAAGTCTCTCAAGAATGATGGCTCACTGGAATGCGTCTTTGACCTGGCCTGG 420
QY 442 GTTCTCGTGAACTTAAACTTTTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC 501
DB 421 GTTCTCGTGAACTTAAACTTTTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC 480
QY 502 GTAAAGCTGTGAGCTGATTTGGAACATGCAAGGTCAATCAATGAGCCTCAAGTCAGTT 561
DB 481 GTAAAGCTGTGAGCTGATTTGGAACATGCAAGGTCAATCAATGAGCCTCAAGTCAGTT 540
QY 562 GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTAGCGGTGGAAGATGGCAACATTATG 621
DB 541 GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTAGCGGTGGAAGATGGCAACATTATG 600
QY 622 GTCTGGATACAGGTGCAACAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT 681
DB 601 GTCTGGATACAGGTGCAACAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT 660
QY 682 GGAGCTCAATATACCTGAGACAGCAAAACCCCTTCAAAACCCAGAGAAACAGAAATCA 741
DB 661 GGAGCTCAATATACCTGAGACAGCAAAACCCCTTCAAAACCCAGAGAAACAGAAATCA 720
QY 742 AAGGACTTGTCTCTCTGCGATTTCCAGCAAAAGTGTACTGTGGTCTCTTTCAAGAC 801
DB 721 AAGGACTTGTCTCTCTCTGCGATTTCCAGCAAAAGTGTACTGTGGTCTCTTTCAAGAC 780
QY 802 GAGAAATCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 861
DB 781 GAGAAATCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 840
QY 862 AAGAAATTAATCTGCTTATCGAACAAGCCATAGATCCAAAGTCTTTCTGTACCAGGT 921
DB 841 AAGAAATTAATCTGCTTATCGAACAAGCCATAGATCCAAAGTCTTTCTGTACCAGGT 900
QY 922 AGCAGCACTCCAAAACCTTGGATATTCAAGTCTGATTTTGGATTTCCAGTGGCTCTACTTTA 981
DB 901 AGCAGCACTCCAAAACCTTGGATATTCAAGTCTGATTTTGGATTTCCAGTGGCTCTACTTTA 960
QY 982 TTTGCTAATTCACAGACGATAACATCTACATGTTTAAATATGACTGGGTGGAAGACTTCT 1041
DB 961 TTTGCTAATTCACAGACGATAACATCTACATGTTTAAATATGACTGGGTGGAAGACTTCT 1020
QY 1042 CCAGTGGCTATTTCAATGGACACAGAACTCTACCTTTTATGTAAATTCAGCCTTAGT 1101
DB 1021 CCAGTGGCTATTTCAATGGACACAGAACTCTACCTTTTATGTAAATTCAGCCTTAGT 1080

QY 1102 CCAGATGACCAAGTCTTTAGTCACTGGCTCAAGTGTAGCTGACCTGCCTACATATGGAAGTC 1161
DB 1081 CCAGATGACCAAGTCTTTAGTCACTGGCTCAAGTGTAGCTGACCTGCCTACATATGGAAGTC 1140
QY 1162 TCCACACCTCGCAACCTCTACTGTGTCTCTGGGTCAATCTCAAGAGGTCAAGTCTCTGTG 1221
DB 1141 TCCACACCTCGCAACCTCTACTGTGTCTCTGGGTCAATCTCAAGAGGTCAAGTCTCTGTG 1200
QY 1222 TGCTGGTTCATCTGACTTCAAAAGATTGCTACCTGTCTGTATGACAATAACATAAAA 1281
DB 1201 TGCTGGTTCATCTGACTTCAAAAGATTGCTACCTGTCTGTATGACAATAACATAAAA 1260
QY 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAACCAAGAGGTGATAAATCTTCCACGGTG 1341
DB 1261 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAACCAAGAGGTGATAAATCTTCCACGGTG 1320
QY 1342 GGTTCGGGCTCTCAGAGAAAAGAGTCAAGACCTGGCTAGTAAACAGTAGTAACGAGTAGC 1401
DB 1321 GGTTCGGGCTCTCAGAGAAAAGAGTCAAGACCTGGCTAGTAAACAGTAGTAACGAGTAGC 1380
QY 1402 CAGAGTACTCTGCGCAAAAGCCCCAGGGTAAAGTGCATTCATCCAAATCTTTCGCCGTCA 1461
DB 1381 CAGAGTACTCTGCGCAAAAGCCCCAGGGTAAAGTGCATTCATCCAAATCTTTCGCCGTCA 1440
QY 1462 TCCGAGCTTGTGCCCAAGCTGTCTGGAGACCTCCCTCTTCCCTCAAACTACTCCTACG 1521
DB 1441 TCCGAGCTTGTGCCCAAGCTGTCTGGAGACCTCCCTCTTCCCTCAAACTACTCCTACG 1500
QY 1522 TTCTCTATTAAACCTCTCTGCCAAGGCCGGTCTCCCATCAACAGAGAGGCTCTCTGC 1581
DB 1501 TTCTCTATTAAACCTCTCTGCCAAGGCCGGTCTCCCATCAACAGAGAGGCTCTCTGC 1560
QY 1582 TCCTCCGTCTCTCCCAAGCACCTTCACTTTTCAAGATGTGATTAAGAACTGGGTGACC 1641
DB 1561 TCCTCCGTCTCTCCCAAGCACCTTCACTTTTCAAGATGTGATTAAGAACTGGGTGACC 1620
QY 1642 CGAACACCTTCTCATCACCACCATCACTCCACCTGCTTCGGAGACCAAGATCATGTCT 1701
DB 1621 CGAACACCTTCTCATCACCACCATCACTCCACCTGCTTCGGAGACCAAGATCATGTCT 1680
QY 1702 CCGAGAAAAGCCCTTATTCTGTGAGCCAGAGTCACTCCCAAGCAGAGGCTTGTCTCGAG 1761
DB 1681 CCGAGAAAAGCCCTTATTCTGTGAGCCAGAGTCACTCCCAAGCAGAGGCTTGTCTCGAG 1740
QY 1762 TCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAAG 1821
DB 1741 TCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAAG 1800
QY 1822 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTGATGGCCCAAGTTGAAATCTTCATTTG 1881
DB 1801 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTGATGGCCCAAGTTGAAATCTTCATTTG 1860
QY 1882 GATCTGTGCTGTGCTGTGTAACAGAGAACCTTAGTAAGGACTCTCTAGGTCTTACC 1941
DB 1861 GATCTGTGCTGTGCTGTGTAACAGAGAACCTTAGTAAGGACTCTCTAGGTCTTACC 1920
QY 1942 AAATCAAGCAAAATGGAAGAGCTGGTACCAGTATCTCAGAGCCTCGGTCTCCTATCAGT 2001
DB 1921 AAATCAAGCAAAATGGAAGAGCTGGTACCAGTATCTCAGAGCCTCGGTCTCCTATCAGT 1980
QY 2002 CCGTATGCTTCAGAAAGCTGTGGAAGCTTACCTCTCTTTTGAGACCTTGTGGAGAGGG 2061
DB 1981 CCGTATGCTTCAGAAAGCTGTGGAAGCTTACCTCTCTTTTGAGACCTTGTGGAGAGGG 2040
QY 2062 TCTGAAATGTGAGGCAAGAGAAATAGTTCGCCAGAGAAATAAAAATCTGGTTGTGGCCATG 2121
DB 2041 TCTGAAATGTGAGGCAAGAGAAATAGTTCGCCAGAGAAATAAAAATCTGGTTGTGGCCATG 2100
QY 2122 GCAGCAAAACGGAAGCTGAGAAATCACTCTCAAGAGTTCGTCTATCCAGACACCCCAAT 2181
DB 2101 GCAGCAAAACGGAAGCTGAGAAATCACTCTCAAGAGTTCGTCTATCCAGACACCCCAAT 2160
QY 2182 TCCAGGAGACAGCGGGAAGACATTGGCAAGCCCGGTCAACCATCAGCCCGAGCTCCATG 2241

DB 2161 TCCAGGAGACAGAGCGGAAGACATTGCCAAGCCCGGTCAACATCAAGCCAGCTCCATG 2220
QY 2242 AGGAAAATCTGCACATATCTTCCATAGAAAAGTCCAGAGGAGCTTCTGTGGTCTGGAACAC 2301
DB 2221 AGGAAAATCTGCACATATCTTCCATAGAAAAGTCCAGAGGAGCTTCTGTGGTCTGGAACAC 2280
QY 2302 TCAACAGAAATATAGATTTCTAACTGAGTGTAGTTACTGAGCTTTGGTCCACTAAACAG 2361
DB 2281 TCAACAGAAATATAGATTTCTAACTGAGTGTAGTTACTGAGCTTTGGTCCACTAAACAG 2340
QY 2362 CTGAGCTTTGGTCCACTAAAACAGATGAAAAATACAAAGAGTACTCTATAACTCTGTGTC 2421
DB 2341 CTGAGCTTTGGTCCACTAAAACAGATGAAAAATACAAAGAGTACTCTATAACTCTGTGTC 2400
QY 2422 TTTAAGAAAGCTGCTTTTTCATTTTATAGACAAAAATCTTTTCAACGCTGAAATGTACCTAA 2481
DB 2401 TTTAAGAAAGCTGCTTTTTCATTTTATAGACAAAAATCTTTTCAACGCTGAAATGTACCTAA 2460
QY 2482 TCTGGTCTACTACCATAAATGTATATGAGCTTCCGAGGATGAATGCTGTGTTAAAT 2541
DB 2461 TCTGGTCTACTACCATAAATGTATATGAGCTTCCGAGGATGAATGCTGTGTTAAAT 2520
QY 2542 TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2601
DB 2521 TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2580
QY 2602 TCATCTCTCTATTAATTAATGATCCAGATTCAGGGGCAAAAACAAAGTTCTTGTGTA 2661
DB 2581 TCATCTCTCTATTAATTAATGATCCAGATTCAGGGGCAAAAACAAAGTTCTTGTGTA 2640
QY 2662 TCCTGAAACTTTCTATGCTCAGTGTGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
DB 2641 TCCTGAAACTTTCTATGCTCAGTGTGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
QY 2722 GGCTGACTGAGAAATCTCTGCTGAGACCCCTGGTCTGTCTGCTCCCAACATGTATA 2781
DB 2701 GGCTGACTGAGAAATCTCTGCTGAGACCCCTGGTCTGTCTGCTCCCAACATGTATA 2760
QY 2782 ATTTTATTTGAAATACATAAATCTTTTCACTATG 2814
DB 2761 ATTTTATTTGAAATACATAAATCTTTTCACTATG 2793

RESULT 11
ADP88308
ID ADP88308 standard; cDNA; 4221 BP.
XX
AC ADP88308;
XX
DT 12-AUG-2004 (first entry)
XX
DE Metastatic breast cancer marker gene RAMP, SEQ ID NO:85.
XX
KW Diagnostic marker; portfolio; diagnostic parameter; gene expression;
KW relative expression; heuristic rule; molecular diagnostic; diagnosis;
KW prognosis; cancer; breast; prostate; ovarian; colorectal; lung;
KW prognostic portfolio; human; metastatic; RAMP; gene; ss.
OS Homo sapiens.
XX
PN EP1349104-A2.
XX
PD 01-OCT-2003.
XX
PF 31-MAR-2003; 2003EP-00252027.
XX
PR 29-MAR-2002; 2002US-0368790P.
XX
PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
PI Jatkoe T;
XX

DR WPI; 2003-815043/77.
XX Selecting a portfolio of markers useful in diagnostic applications
PT comprises defining diagnostic parameters and establishing a relationship
PT among the diagnostic parameters that identifies an optimized portfolio of
XX markers.
XX Example 3; SEQ ID NO 85; 154pp; English.
XX The invention relates to a method of selecting a portfolio of markers
CC (e.g., genes) for use in a diagnostic application. The method involves:
CC (a) defining diagnostic parameters; (b) establishing a relationship among
CC the diagnostic parameters so that they are optimised; and (c) selecting
CC the optimal group of markers (the portfolio) for the diagnostic
CC application. The diagnostic parameters can include a measure of the
CC relative degree of expression of a gene, a measure of the variation in
CC the measurement of the degree of gene expression, and the relationship
CC between the diagnostic and discriminating parameters can be a mean
CC variance relationship. The method further comprises the application of a
CC heuristic rule. The invention also relates to a general-purpose computer
CC programmed to identify a portfolio of markers, instructions for
CC performing the method of the invention, and a diagnostic portfolio
CC comprising genes selected according to the method. The method of the
CC invention can be used in conjunction with any method for determining the
CC gene expression of relevant cells as well as protein-based methods of
CC determining gene expression. The method of the invention can be used in
CC screening, diagnosing and prognosing diseases associated with changes in
CC the expression of multiple genes (e.g., breast, prostate, ovarian,
CC colorectal or lung cancer). By grouping these markers into portfolios,
CC the most reliable results can be obtained with the smallest number of
CC markers necessary to obtain such a result, which is particularly
CC important in multiple-step assays such as nucleic amplification methods.
CC The present sequence represents a human diagnostic marker gene that is a
CC member of an optimised portfolio of 28 genes for prognosing breast
XX cancer.
XX Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;
Query Match 98.5%; Score 2789.8; DB 11; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 22 CGATAACGATTGTGTGTGAGAGGCGCAACGTGCGATTTCGTGAACTCTGGAGGCATT 81
DB 1 CGATAACGATTGTGTGTGAGAGGCGCAACGTGCGATTTCGTGAACTCTGGAGGCATT 60
QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTGCGGTG 141
DB 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTGCGGTG 120
QY 142 CTCGCGCAGCCAGCTTGGGCTCTGAGAAATGGAATGGTCTTTCACATACCTCTTCAA 201
DB 121 CTCGCGCAGCCAGCTTGGGCTCTGAGAAATGGAATGGTCTTTCACATACCTCTTCAA 180
QY 202 TCCCTTCTCAGCTGTTATCAGTGCAGTGTGTAATGATGAACACACTTCTTATGAGAAACA 261
DB 181 TCCCTTCTCAGCTGTTATCAGTGCAGTGTGTAATGATGAACACACTTCTTATGAGAAACA 240
QY 262 GGAGTCCCAAGTTCCTCTTGGATGATACCTTCTCTTCTGCTCCCAATATGAAACATGTA 321
DB 241 GGAGTCCCAAGTTCCTCTTGGATGATACCTTCTCTTCTGCTCCCAATATGAAACATGTA 300
QY 322 CTAGCAGTTGCCAATGAAGAAGCTTTGTGATGATGATGATGATGATGATGATGATGATGAT 381
DB 301 CTAGCAGTTGCCAATGAAGAAGCTTTGTGATGATGATGATGATGATGATGATGATGATGAT 360
QY 382 AGAAGAAGTGTCTCAAGAATGGATGGCTCACTGGAATGCCGTCTTTGACCTGGCTGG 441
DB 361 AGAAGAAGTGTCTCAAGAATGGATGGCTCACTGGAATGCCGTCTTTGACCTGGCTGG 420
QY 442 GTTCTGTGTGAATTAACCTTGTACAGCAGCAGGCTGATCAAAACGCCAAATTTTGGGAC 501
DB 421 GTTCTGTGTGAATTAACCTTGTACAGCAGCAGGCTGATCAAAACGCCAAATTTTGGGAC 480

QY 502 GTAAAAGCTGTGAGCTGATTGGAAACATGCAAAAGCTCATCAATGCAGCCTCAAGTCAGTT 561
DB 481 GTAAAAGCTGTGAGCTGATTGGAAACATGCAAAAGCTCATCAATGCAGCCTCAAGTCAGTT 540
QY 562 GCCTTTTCTAAAGTTTGAGAAAGCTGTATTCTGTGTCGGGTGAAAGAGATGGCAACATTATG 621
DB 541 GCCTTTTCTAAAGTTTGAGAAAGCTGTATTCTGTGTCGGGTGAAAGAGATGGCAACATTATG 600
QY 622 GTCTGGGATACAGGTGCAACAAAAGATGGGTTTTATAGCAGTCAATCAAAATCAGT 681
DB 601 GTCTGGGATACAGGTGCAACAAAAGATGGGTTTTATAGCAGTCAATCAAAATCAGT 660
QY 682 GGAGCTCAAAATACCTCAGACAAAGCAACCCCTTCAAAACCCCAAGAAAGAAACAGAAATCA 741
DB 661 GGAGCTCAAAATACCTCAGACAAAGCAACCCCTTCAAAACCCCAAGAAAGAAACAGAAATCA 720
QY 742 AAAGSACTTGTCTCTCTGTGATTTCCAGCAAAAGTGTATCTGTGGTCTCTTTTCAAGAC 801
DB 721 AAAGSACTTGTCTCTCTGTGATTTCCAGCAAAAGTGTATCTGTGGTCTCTTTTCAAGAC 780
QY 802 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGATTTAGT 861
DB 781 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGATTTAGT 840
QY 862 AAGATATATACCTGCTTATCGACAAAGAACCCATAGCATCCAAAGTCTTTCTGTACCAGCT 921
DB 841 AAGATATATACCTGCTTATCGACAAAGAACCCATAGCATCCAAAGTCTTTCTGTACCAGCT 900
QY 922 AGCAGCACTCGAAAACCTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
DB 901 AGCAGCACTCGAAAACCTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
QY 982 TTTGCTAATTGCAACAGCAGATAACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT 1041
DB 961 TTTGCTAATTGCAACAGCAGATAACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT 1020
QY 1042 CCAGTGGCTATTTTCAATGGACACCAAGAACTCTACTTTTATGTAAATTCAGCCTTAGT 1101
DB 1021 CCAGTGGCTATTTTCAATGGACACCAAGAACTCTACTTTTATGTAAATTCAGCCTTAGT 1080
QY 1102 CCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAATGAGCTGCCTACATATGGAAGTCT 1161
DB 1081 CCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAATGAGCTGCCTACATATGGAAGTCT 1140
QY 1162 TCCACACCTGGCAACCTCTACTGTCTCTGGTCAATCTCAGAGGTCACGTCTGTG 1221
DB 1141 TCCACACCTGGCAACCTCTACTGTCTCTGGTCAATCTCAGAGGTCACGTCTGTG 1200
QY 1222 TGCTGGTGTCCATCTGACTTCCAAAAGATTCTACCTGTTCTGTGACAAATACACTAAAA 1281
DB 1201 TGCTGGTGTCCATCTGACTTCCAAAAGATTCTACCTGTTCTGTGACAAATACACTAAAA 1260
QY 1282 ATCTGGCCCTTTGAATFAGAGGCTTAGAGGAGAAACCCAGAGGTGATAAACTTTTCCACGGTG 1341
DB 1261 ATCTGGCCCTTTGAATFAGAGGCTTAGAGGAGAAACCCAGAGGTGATAAACTTTTCCACGGTG 1320
QY 1342 GGTGGGGCTCTCAGAGAAAGAGAGTCAAGACCTGGCTAGTAAACAGTAACGAGTAGC 1401
DB 1321 GGTGGGGCTCTCAGAGAAAGAGAGTCAAGACCTGGCTAGTAAACAGTAACGAGTAGC 1380
QY 1402 CAGAGTACTCTGCGAAAGCCCGAGGGTAAAGTGAATCCATCCAAATCTTCCCGGTCA 1461
DB 1381 CAGAGTACTCTGCGAAAGCCCGAGGGTAAAGTGAATCCATCCAAATCTTCCCGGTCA 1440
QY 1462 TCCGAGCTGTGCCCCCAAGCTGTGTCGAGACCTCTCTCTTCTTCAAACTACTCTTACG 1521
DB 1441 TCCGAGCTGTGCCCCCAAGCTGTGTCGAGACCTCTCTCTTCTTCAAACTACTCTTACG 1500
QY 1522 TTCTCTATTAAAACTCTCTCTGCCAAGCCCGGTCTCCCATCAACAGAAAGGCTCTGTCT 1581
DB 1501 TTCTCTATTAAAACTCTCTCTGCCAAGCCCGGTCTCCCATCAACAGAAAGGCTCTGTCT 1560

QY 1582 TCCTCGGTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTCGATTAGAACTGGGTGACC 1641
DB 1561 TCCTCGGTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTCGATTAGAACTGGGTGACC 1620
QY 1642 CGAACAACCTTCCTCATCACCAACCATCACTCCACCTGCTTCGGAGACCAAGATCATGTCT 1701
DB 1621 CGAACAACCTTCCTCATCACCAACCATCACTCCACCTGCTTCGGAGACCAAGATCATGTCT 1680
QY 1702 CCGAGAAAAGCCCTTATTCCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGCTCTGAG 1761
DB 1681 CCGAGAAAAGCCCTTATTCCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGCTCTGAG 1740
QY 1762 TCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAAG 1821
DB 1741 TCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAAG 1800
QY 1822 TGTGTGAGAGTGTGTAACCTGTGTGACTGAGCTTGATGGCCAAAGTTGAAAATCTTCATTTTG 1881
DB 1801 TGTGTGAGAGTGTGTAACCTGTGTGACTGAGCTTGATGGCCAAAGTTGAAAATCTTCATTTTG 1860
QY 1882 GATCTGTGCTGCTTGTCTGCTGAACCAAGGACCTTGTAGTAAGACTCTCTAGGTCCTACC 1941
DB 1861 GATCTGTGCTGCTTGTCTGCTGAACCAAGGACCTTGTAGTAAGACTCTCTAGGTCCTACC 1920
QY 1942 AAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCAGAGCCTCCGTCCTCTATCAGT 2001
DB 1921 AAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCAGAGCCTCCGTCCTCTATCAGT 1980
QY 2002 CCGTATGCTTCAGAAAGCTGTGGAAGCTACCTCTTCCTTTGAGAGCTTTGTGGAGAGGG 2061
DB 1981 CCGTATGCTTCAGAAAGCTGTGGAAGCTACCTCTTCCTTTGAGAGCTTTGTGGAGAGGG 2040
QY 2062 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACTGGTTGTGGCCATG 2121
DB 2041 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACTGGTTGTGGCCATG 2100
QY 2122 GCAGCCAAACGGAAGGCTGAGAAATCAATCTCCAGAGTCCGTCAATCCAGACACCCCAAT 2181
DB 2101 GCAGCCAAACGGAAGGCTGAGAAATCAATCTCCAGAGTCCGTCAATCCAGACACCCCAAT 2160
QY 2182 TCCAGGAGACAGCGGGAAGACATTTGCCAAGCCCGGTCCACCATCAGCCAGCTCCCATG 2241
DB 2161 TCAGAGAGACAGCGGGAAGACATTTGCCAAGCCCGGTCCACCATCAGCCAGCTCCCATG 2220
QY 2242 AGGAAATCTGCATACACTTCCATAGAAAGTCCAGAGGAGCTTCTGTGGTCCCTGAACAC 2301
DB 2221 AGGAAATCTGCATACACTTCCATAGAAAGTCCAGAGGAGCTTCTGTGGTCCCTGAACAC 2280
QY 2302 TCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTTAAACAAAG 2361
DB 2281 TCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTTAAACAAAG 2340
QY 2362 CTGAGCTTTGGTCCACTTAAACAGATGAAATAACAAGAGTACTCTATATCTCTGGTC 2421
DB 2341 CTGAGCTTTGGTCCACTTAAACAGATGAAATAACAAGAGTACTCTATATCTCTGGTC 2400
QY 2422 TTTTAAAGAAAGCTGCCCTTTTCATTTTATAGACAAAATCTTTTCAACGCTGAATGTACCTAA 2481
DB 2401 TTTTAAAGAAAGCTGCCCTTTTCATTTTATAGACAAAATCTTTTCAACGCTGAATGTACCTAA 2460
QY 2482 TCTGGTTCTTACTACCMATATGATATGAGCTTCCGAGGATGAATGCTGTGTTTAAATTT 2541
DB 2461 TCTGGTTCTTACTACCMATATGATATGAGCTTCCGAGGATGAATGCTGTGTTTAAATTT 2520
QY 2542 TCATAAAGTAAATTTGTCTACTCTAGCATTTTCAATGAAATGAGTCTTCACCTTTTAAATAT 2601
DB 2521 TCATAAAGTAAATTTGTCTACTCTAGCATTTTCAATGAAATGAGTCTTCACCTTTTAAATAT 2580
QY 2602 TCATCTTCTCTATAATATGACATCCAGTTCATGGAGGCAAAAACAAAGTTCTTGTTA 2661
DB 2581 TCATCTTCTCTATAATATGACATCCAGTTCATGGAGGCAAAAACAAAGTTCTTGTTA 2640
QY 2662 TCTTGAAACTTCTATGCTCAGTGGGAAAGTATCTGCCAGCCACAGCATGAGCCCTGTGAA 2721

DB 2641 TCCTGAACAATTTCTATGCTCAGTGGAAAGATATCTGCCAGCCACAGCATGAGCCCTGTGAA 2700
QY 2722 GGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTCTGTCTGCTTCCCTCCAAACATGTATA 2781
DB 2701 GGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTCTGTCTGCTTCCCTCCAAACATGTATA 2760
QY 2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
DB 2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793

RESULT 12
ADP07299 standard; DNA; 4221 BP.
XX
AC ADP07299;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human L2DTL DNA.
XX
KW ds; proliferative disease; breast; methylation; CpG; bisulfite; human.
XX
OS Homo sapiens.
XX
PN DE10255104-A1.
XX
PD 11-MAR-2004.
XX
PF 26-NOV-2002; 2002DE-01055104.
XX
PR 27-AUG-2002; 2002DE-01039313.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Maier S;
XX
DR WPI; 2004-284340/27.
DR GENBANK; NM_016448.
XX
PT Analyzing proliferative diseases of breast cells, useful e.g. for
PT diagnosis, prognosis and treatment of breast cancer, by determining
PT methylation status of specific genes.
XX
PS Claim 1; Page; 22pp; German.
XX
CC This invention describes a novel method of analysing proliferative
CC diseases of breast cells by determining the methylation status of certain
CC genes. The invention also describes nucleic acids or their complements,
CC oligomers, especially oligonucleotides or peptide nucleic acid oligomers,
CC that hybridise to, or are identical with, any of the nucleic acids, the
CC preparation of an oligomer array for analysing proliferative diseases of
CC breast cells that are associated with the methylation status of CpG
CC dinucleotides of the genes by bonding at least one oligomer to a solid
CC phase and a kit comprising a bisulfite reagent and the oligomer. The
CC nucleic acids are genomic sequences (5' - and/or regulatory and/or CpG-
CC rich regions). The base sequence of the oligomer includes at least one
CC CpG island, especially with C in the middle third of the sequence. The
CC process involves a genomic DNA sample treated chemically, specifically
CC with a bisulfite reagent, to convert non-methylated C to uracil or some
CC other base having base-pairing properties different from those of C.
CC Fragments of the treated DNA are then amplified, using primers and a
CC polymerase and the methylation status of the genomic CpG dinucleotides is
CC determined by analysis of the amplicons, particularly by hybridisation to
CC the oligomer, optionally with extension of the hybridised oligomer by at
CC least one base, or detection is by sequencing. The amplification may use
CC methylation-specific primers. Alternatively, genomic DNA is extracted
CC from a sample and digested with methylation-specific restriction enzymes,
CC then the digestion fragments detected, optionally after amplification. In
CC either method, more than 10 fragments of 100-200 bp are amplified in a
CC single reaction vessel, using a heat-resistant DNA polymerase in PCR. The
CC amplicons carry detectable markers, e.g. fluorophores, radioisotopes

CC and/or releasable fragments of known mass that can be detected by mass
CC spectrometry. The method is used for characterisation, classification,
CC differentiation, staging, phase-estimation, diagnosis and/or therapy of
CC proliferative diseases of breast cells. The method provides very specific
CC classification of proliferative diseases, allowing better treatment. It
CC can both characterise methylation status and detect single-nucleotide
CC polymorphisms. This sequence represents human gene used to illustrate the
CC method of the invention. NOTE: This sequence does not appear in the
CC printed specification but has been retrieved from Genbank.
XX
SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;
Query Match 98.5%; Score 2789.8; DB 12; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 22 CGATAACGATTGTGTGTGAGAGGCGCAACGTCGATTTCTGCTGAACTTGGAGGCATT 81
DB 1 CGATAACGATTGTGTGTGAGAGGCGCAACGTCGATTTCTGCTGAACTTGGAGGCATT 60
QY 82 TCTACGACTTTTCTCTCAGTCGAGGCTTTTCTCCGACCTGATGCTCTTCAATTGGTG 141
DB 61 TCTACGACTTTTCTCTCAGTCGAGGCTTTTCTCCGACCTGATGCTCTTCAATTGGTG 120
QY 142 CTCGCGCAGCCGAGCTTGGCGTCTCGAGAAATGATGGTCTTCAATACCCCTCTTCAA 201
DB 121 CTCGCGCAGCCGAGCTTGGCGTCTGAGAAATGATGGTCTTCAATACCCCTCTTCAA 180
QY 202 TCCCTTCTGACTGGTTATCAGTCAGTCAGTGAATGATGAACACACTTCTTATGGAGAAACA 261
DB 181 TCCCTTCTGACTGGTTATCAGTCAGTCAGTGAATGATGAACACACTTCTTATGGAGAAACA 240
QY 262 GGNATCCGAGTTCTCTTTTGGATGATCTTCTTCTGCTCCCAATATGGAACATGTA 321
DB 241 GGNATCCGAGTTCTCTTTTGGATGATCTTCTTCTGCTCCCAATATGGAACATGTA 300
QY 322 CTAGCAGTTGCCAATGAAGAGCTTTGTTGCATGTTATAACACAGAAATCAACAAGTTTC 381
DB 301 CTAGCAGTTGCCAATGAAGAGCTTTGTTGCATGTTATAACACAGAAATCAACAAGTTTC 360
QY 382 AGAAGAGTGTCTTCAAGAAATGATGGTCTCACTGGAATGCCGTCTTTGACCTGGCCTGG 441
DB 361 AGAAGAGTGTCTTCAAGAAATGATGGTCTCACTGGAATGCCGTCTTTGACCTGGCCTGG 420
QY 442 GTTCTGTTGAACTTAACTTGTATACAGCAGCAGTGTATCAACAGCCCAAAATTTGGGAC 501
DB 421 GTTCTGTTGAACTTAACTTGTATACAGCAGCAGTGTATCAACAGCCCAAAATTTGGGAC 480
QY 502 GTAAAAGCTGGTGAGCTGATTGGAAACATGCAAGAGTCAATGACAGCCTCAAGTCAGTT 561
DB 481 GTAAAAGCTGGTGAGCTGATTGGAAACATGCAAGAGTCAATGACAGCCTCAAGTCAGTT 540
QY 562 GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG 621
DB 541 GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG 600
QY 622 GTCTGGGATACAGGTGCAACAAAAGATGGTTTTATAGGCAAGTGAATCAATACGT 681
DB 601 GTCTGGGATACAGGTGCAACAAAAGATGGTTTTATAGGCAAGTGAATCAATACGT 660
QY 682 GGAGCTCACAACTCTCAGACAGCAAAACCCCTTCAAAAACCCCAAGAAAGAAACAGAAATCA 741
DB 661 GGAGCTCACAACTCTCAGACAGCAAAACCCCTTCAAAAACCCCAAGAAAGAAACAGAAATCA 720
QY 742 AAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTTTTCAAGAC 801
DB 721 AAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTTTTCAAGAC 780
QY 802 GAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGAATCAAGTATGGGATTTACGT 861
DB 781 GAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATGAATCAAGTATGGGATTTACGT 840
QY 862 AAGAAATTATCTGCTTATCGACAAGAACCCATAGCATCCAGTCTTCTGTTACCCAGGT 921

DB 841 AAGAATTATCTGCTTATCGACAAGAACCCATAGCATCCAAGTCTTTCTCTGATCCAGGT 900
QY 922 AGCAGACTCGAAACTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTTA 981
DB 901 AGCAGACTCGAAACTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTTA 960
QY 982 TTTGCTAAATGACAGACGATAAATCTACATCTGTTTAAATATGACTGGGTTGAAGACTTCT 1041
DB 961 TTTGCTAAATGACAGACGATAAATCTACATCTGTTTAAATATGACTGGGTTGAAGACTTCT 1020
QY 1042 CCAGTGGCTATTTTCAATGGACACAGAACTCTACTCTTTTATGTAAAAATCCAGCCTTAGT 1101
DB 1021 CCAGTGGCTATTTTCAATGGACACAGAACTCTACTCTTTTATGTAAAAATCCAGCCTTAGT 1080
QY 1102 CCAGATGACAGTCTTTTGTAGTCAAGTGAAGCTGATGATGATGATGATGATGATGATGATG 1161
DB 1081 CCAGATGACAGTCTTTTGTAGTCAAGTGAAGCTGATGATGATGATGATGATGATGATGATG 1140
QY 1162 TCCACACCTGGCAACCTCTACTGTGCTCTCTGGGTCAATCTCTCAAGAGGTCAAGTCTGTG 1221
DB 1141 TCCACACCTGGCAACCTCTACTGTGCTCTCTGGGTCAATCTCTCAAGAGGTCAAGTCTGTG 1200
QY 1222 TGCTGGTGTCCATCTGACTTCAAAAAGATTGCTACTGTTCTGTGATGACAAATACATAAAA 1281
DB 1201 TGCTGGTGTCCATCTGACTTCAAAAAGATTGCTACTGTTCTGTGATGACAAATACATAAAA 1260
QY 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCAGGGTG 1341
DB 1261 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCAGGGTG 1320
QY 1342 GGTTCGGGCTCTCAGAGAAAGAGGTCAAGACCTGGCTAGTAACAGTAAACAGAGTAGC 1401
DB 1321 GGTTCGGGCTCTCAGAGAAAGAGGTCAAGACCTGGCTAGTAACAGTAAACAGAGTAGC 1380
QY 1402 CAGAGTACTCTCTGCAAGAGCCCAAGGGTAAAGTCAATCCAATCTTTCCCGGTCA 1461
DB 1381 CAGAGTACTCTCTGCAAGAGCCCAAGGGTAAAGTCAATCCAATCTTTCCCGGTCA 1440
QY 1462 TCCGAGCTTGTGCCCCAAGCTGTCTGGAGACCTCCCTCTTCCCTTCAAAATCTCTAGG 1521
DB 1441 TCCGAGCTTGTGCCCCAAGCTGTCTGGAGACCTCCCTCTTCCCTTCAAAATCTCTAGG 1500
QY 1522 TTTCTATTAAAACTCTCTGCAAGGCCGCTCCCATCAACAGAGAGGCTCTGTG 1581
DB 1501 TTTCTATTAAAACTCTCTGCAAGGCCGCTCCCATCAACAGAGAGGCTCTGTG 1560
QY 1582 TCCTCCGCTCTCTCCCAAGCCACCTTTCATCTTTCAAGATGTGATTTAGAAACTGGGGTACC 1641
DB 1561 TCCTCCGCTCTCTCCCAAGCCACCTTTCATCTTTCAAGATGTGATTTAGAAACTGGGGTACC 1620
QY 1642 CGAAACCTTCTCTATACACCCCATCTCACTGCTGTGGAGACCAAGATCATGTCT 1701
DB 1621 CGAAACCTTCTCTATACACCCCATCTCACTGCTGTGGAGACCAAGATCATGTCT 1680
QY 1702 CCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1761
DB 1681 CCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1740
QY 1762 TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1821
DB 1741 TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1800
QY 1822 TGTGTGAAGAGTTGAACTGTGTGACTGAGCTTGTATGGCAAGTTGAAATCTTCTATTG 1881
DB 1801 TGTGTGAAGAGTTGAACTGTGTGACTGAGCTTGTATGGCAAGTTGAAATCTTCTATTG 1860
QY 1882 GATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1941
DB 1861 GATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
QY 1942 AAATCAAGCAAAATTTGAAGGAGCTGGTATCTCAGAGCTTCCGCTCTCTCTATCAGT 2001

QY 382 AGAAGAAGTGCTTCAAAGAATGGATGGCTCACTGGAATGCCGTCTTTTGACCTGGCGTGG 441
DB |||||
QY 361 AGAAGAAGTGCTTCAAAGATGGATGGCTCACTGGAATGCCGTCTTTTGACCTGGCGTGG 420
DB |||||
QY 442 GTTCTCGGTGAACTTAAACTTGTGTTACAGCAGCAGGTGATCAAAACAGCCAAAATTTGGGAC 501
DB |||||
QY 421 GTTCTCGGTGAACTTAAACTTGTGTTACAGCAGCAGGTGATCAAAACAGCCAAAATTTGGGAC 480
DB |||||
QY 502 GTAAGAAGCTGAGCTGATGGAACATGCAGAAGTTCATCAATGCAGAGCTCAAGTCAGTT 561
DB |||||
QY 481 GTAAAGAGCTGGTGAGCTGATGGAAATGCAAAAGGTCATCAATGCAAGCTCAAGTCAGTT 540
DB |||||
QY 562 GCCTTTTCTAAAGTTTGAGAAAGCTGTATCTGTACGGGTGGAAAGAGATGGCAACATTATG 621
DB |||||
QY 541 GCCTTTTCTAAAGTTTGAGAAAGCTGTATCTGTACGGGTGGAAAGAGATGGCAACATTATG 600
DB |||||
QY 622 GTCTGGGATACAGGTGCAAAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAAT 681
DB |||||
QY 601 GTCTGGGATACAGGTGCAAAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAAT 660
DB |||||
QY 682 GGAAGTCAAAATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATCA 741
DB |||||
QY 661 GGAAGTCAAAATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATCA 720
DB |||||
QY 742 AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTCTTTCAAGAC 801
DB |||||
QY 721 AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTCTTTCAAGAC 780
DB |||||
QY 802 GAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAGTATGGGATTTACGT 861
DB |||||
QY 781 GAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAGTATGGGATTTACGT 840
DB |||||
QY 862 AAGAATTAATCTCTTATCGACAAGAAACCCATAGCATCCAAGTCTTTCTGTACCCAGT 921
DB |||||
QY 841 AAGAATTAATCTCTTATCGACAAGAAACCCATAGCATCCAAGTCTTTCTGTACCCAGT 900
DB |||||
QY 922 AGCAGCACTCGAAACCTGGATATCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
DB |||||
QY 901 AGCAGCACTCGAAACCTGGATATCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
DB |||||
QY 982 TTTGCTAAATGGCAGACGATAAACATCTACATGTTTTAAATATGACTGGGTTGAAGACTTCT 1041
DB |||||
QY 961 TTTGCTAAATGGCAGACGATAAACATCTACATGTTTTAAATATGACTGGGTTGAAGACTTCT 1020
DB |||||
QY 1042 CAGTGGCTATTTTCAATGGACACCAAGACTACCTCTTTTATGTAAATTCAGCCTTAGT 1101
DB |||||
QY 1021 CCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTATGTAAATTCAGCCTTAGT 1080
DB |||||
QY 1102 CCAGATGACAGTTTTTATGTAGTGCTCAAGTGATGAAGCTGCCTACATATGGAGGTC 1161
DB |||||
QY 1081 CCAGATGACAGTTTTTATGTAGTGCTCAAGTGATGAAGCTGCCTACATATGGAGGTC 1140
DB |||||
QY 1162 TCCACACCTCGCAACCTCTACTGTGCTCTGGGTCAATCTCAAGAGGTCAAGTCTGTG 1221
DB |||||
QY 1141 TCCACACCTCGCAACCTCTACTGTGCTCTGGGTCAATCTCAAGAGGTCAAGTCTGTG 1200
DB |||||
QY 1222 TGTGTGTGCTCAATCTGACTTCAAAAGATTGCTPACTGTGTTGATGACAATACACTAAAA 1281
DB |||||
QY 1201 TGTGTGTGCTCAATCTGACTTCAAAAGATTGCTPACTGTGTTGATGACAATACACTAAAA 1260
DB |||||
QY 1282 ATCTGGCGCTTGAATAGGCTTAGAGGAGAACACAGAGGTGATTAACCTTTCCAGGTC 1341
DB |||||
QY 1261 ATCTGGCGCTTGAATAGGCTTAGAGGAGAACACAGAGGTGATTAACCTTTCCAGGTC 1320
DB |||||
QY 1342 GGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACCTGGCTAGTAAACAGTAAACAGGTAGC 1401
DB |||||
QY 1321 GGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACCTGGCTAGTAAACAGTAAACAGGTAGC 1380
DB |||||
QY 1402 CAGAGTACTCTCTGCCAAAGCCCGGTTAAAGTGAATCCATCCAAATTTCTTCCCGCTCA 1461
DB |||||
QY 1381 CAGAGTACTCTCTGCCAAAGCCCGGTTAAAGTGAATCCATCCAAATTTCTTCCCGCTCA 1440
DB |||||

QY 1462 TCCGAGAGCTTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATCTCTACG 1521
DB |||||
QY 1441 TCCGAGAGCTTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATCTCTACG 1500
DB |||||
QY 1522 TTTCTCTATTAAAAACCTCTCTGCGCAAGGCCCGGTCTCCCATCAACAGAAAGAGGCTCTGTCT 1581
DB |||||
QY 1501 TTTCTCTATTAAAAACCTCTCTGCGCAAGGCCCGGTCTCCCATCAACAGAAAGAGGCTCTGTCT 1560
DB |||||
QY 1582 TCCTCGGTCTCTCCCAAGCCACCTTTCATCTTTCAAGATGTCTGATTAGAAACTGGGTGACC 1641
DB |||||
QY 1561 TCCTCGGTCTCTCCCAAGCCACCTTTCATCTTTCAAGATGTCTGATTAGAAACTGGGTGACC 1620
DB |||||
QY 1642 CGAACACCTTCTCTATCACTACCCACCTCTCCACCTGTCTCGGAGACCAAGATCATGTCT 1701
DB |||||
QY 1621 CGAACACCTTCTCTATCACTACCCACCTCTCCACCTGTCTCGGAGACCAAGATCATGTCT 1680
DB |||||
QY 1702 CGGAGAAAGCCCTTATTTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1761
DB |||||
QY 1681 CGGAGAAAGCCCTTATTTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1740
DB |||||
QY 1762 TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAG 1821
DB |||||
QY 1741 TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAG 1800
DB |||||
QY 1822 TGTGTAAAGAGTTGAACTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAAATCTTCAITTTG 1881
DB |||||
QY 1801 TGTGTAAAGAGTTGAACTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAAATCTTCAITTTG 1860
DB |||||
QY 1882 GATCTGTGCTGCTTGTGTGTAAACAGGAAGACCTTTAGTAAGGACTCTCTAGGTCTTACC 1941
DB |||||
QY 1861 GATCTGTGCTGCTTGTGTGTAAACAGGAAGACCTTTAGTAAGGACTCTCTAGGTCTTACC 1920
DB |||||
QY 1942 AAATCAAGCAAAATTTGAAGGAGCTGTGAAACGCTTCTCTTTGAGACCTTTGTGGAGAGG 2001
DB |||||
QY 1921 AAATCAAGCAAAATTTGAAGGAGCTGTGAAACGCTTCTCTTTGAGACCTTTGTGGAGAGG 1980
DB |||||
QY 2002 CCGTATGCTTTCAGAAAGCTGTGAAACGCTTCTCTTTGAGACCTTTGTGGAGAGG 2061
DB |||||
QY 1981 CCGTATGCTTTCAGAAAGCTGTGAAACGCTTCTCTTTGAGACCTTTGTGGAGAGG 2040
DB |||||
QY 2062 TCTGAAATGTGATGAGCAAAAGATAAGTTTCCCAGAGATAAATACTGGTTGTGGCCATG 2121
DB |||||
QY 2041 TCTGAAATGTGATGAGCAAAAGATAAGTTTCCCAGAGATAAATACTGGTTGTGGCCATG 2100
DB |||||
QY 2122 GCAGCAAAACGGAAGGCTGGAATCCATCTCCAGAACTCCGTCATCCAGACACCCCAAT 2181
DB |||||
QY 2101 GCAGCAAAACGGAAGGCTGGAATCCATCTCCAGAACTCCGTCATCCAGACACCCCAAT 2160
DB |||||
QY 2182 TCCAGGAGACAGCGGGAAGACATTGCAAGCCCGGTCAACCATCAAGCCCAAGCTCCCATG 2241
DB |||||
QY 2161 TCCAGGAGACAGCGGGAAGACATTGCAAGCCCGGTCAACCATCAAGCCCAAGCTCCCATG 2220
DB |||||
QY 2242 AGGAAATCTGCACATCTTCCATAGAAAGTCCCAGGAGGACTTCTGTGGTCTCTGAAACAC 2301
DB |||||
QY 2221 AGGAAATCTGCACATCTTCCATAGAAAGTCCCAGGAGGACTTCTGTGGTCTCTGAAACAC 2280
DB |||||
QY 2302 TCACAGAAATTTAGATTTCTAATCTGAGTGAGTACTGAGCTTTGGTCCACTAAACACAG 2361
DB |||||
QY 2281 TCACAGAAATTTAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACACAG 2340
DB |||||
QY 2362 CTGAGCTTTGGTCCACTAAAAACAAGATGAAAAATACAAGAGTGACTCTATTAATCTGGTC 2421
DB |||||
QY 2341 CTGAGCTTTGGTCCACTAAAAACAAGATGAAAAATACAAGAGTGACTCTATTAATCTGGTC 2400
DB |||||
QY 2422 TTTTAAAGAAAGCTGCTTTTTCATTTTTTAGACAAAAATCTTTTCAACGCTGAAATGTACCTAA 2481
DB |||||
QY 2401 TTTTAAAGAAAGCTGCTTTTTCATTTTTTAGACAAAAATCTTTTCAACGCTGAAATGTACCTAA 2460
DB |||||
QY 2482 TCTGTGTTCTATACATTAATGTATATGACGCTTCCGAGGATGAATGTCTGTGTTTAAATTT 2541
DB |||||
QY 2461 TCTGTGTTCTATACATTAATGTATATGACGCTTCCGAGGATGAATGTCTGTGTTTAAATTT 2520
DB |||||
QY 2542 TCATAAAGTAAATTTTGTCTACCTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAAAATTAT 2601
DB |||||

Db 2521 TCATAAGTAATTTGTCTACTCTAGCAATTTGAATGAATAGCTTCACACTTTTAAATAT 2580
Qy 2602 TCATCTTCTCTATAATAATAGACATCCAGTTCATGGAGGCAAAAACAAGTTTCTTGTTA 2661
Db 2581 TCATCTTCTCTATAATAATAGACATCCAGTTCATGGAGGCAAAAACAAGTTTCTTGTTA 2640
Qy 2662 TCCTGAACTTTCTATGCTCACTGGAAGTATCTGCCAGCCACAGCATGAGGCCCTGTGAA 2721
Db 2641 TCCTGAACTTTCTATGCTCACTGGAAGTATCTGCCAGCCACAGCATGAGGCCCTGTGAA 2700
Qy 2722 GCGTCACTGAGAAATCTCTGCTGAGACCCCTGTTCTGTTCTGCCCAACATGTATA 2781
Db 2701 GCGTCACTGAGAAATCTCTGCTGAGACCCCTGTTCTGTTCTGCCCAACATGTATA 2760
Qy 2782 ATTTTATTTGAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAATACATAATCTTTTCACTATG 2793
RESULT 14
ID ADO20231
XX ADO20231 standard; cDNA; 4221 BP.
AC ADO20231;
XX ADO20231;
DT 12-AUG-2004 (first entry)
XX Human PRO polynucleotide #567.
DE
XX Human; PRO; gene; ss; immune related disorder;
KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
KW renal disease; demyelinating disease; central nervous system;
KW peripheral nervous system; demyelinating polyneuropathy;
KW Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX Homo sapiens.
OS
XX WO2004043361-A2.
PN
XX 27-MAY-2004.
PD
XX
PF 06-NOV-2003; 2003WO-US035268.
PP
XX 08-NOV-2002; 2002US-0425235P.
PR (GETH) GENENTECH INC.
XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
PI
XX WPI; 2004-420067/39.
DR P-PSDB; ADO20232.
DR
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthropathy.
XX
PS Claim 1; SEQ ID NO 1202; 1731pp; English.
XX
CC The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the

CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polynucleotide of
CC the invention.
XX
SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;
Query Match 98.5%; Score 2789.8; DB 12; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 22 CGATAACGATTGTGTGAGAGCGCAACGCTGCGATTTCTGCTGAACCTTTGGAGGCATT 81
Db 1 CGATAACGATTGTGTGAGAGCGCAACGCTGCGATTTCTGCTGAACCTTTGGAGGCATT 60
Qy 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCGACCCCTGATGCTCTTCAATTCGGTG 141
Db 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCGACCCCTGATGCTCTTCAATTCGGTG 120
Qy 142 CTCGCCAGCCCGAGCTTGGGCTCTGAGAAATGAGTGGTCTTACAAATACCTCTTCAA 201
Db 121 CTCGCCAGCCCGAGCTTGGGCTCTGAGAAATGAGTGGTCTTACAAATACCTCTTCAA 180
Qy 202 TCCCTTCTGACTGTTATCAGTGCAGTGGTAATGATGAACACACTTCTTATGAGAAACA 261
Db 181 TCCCTTCTGACTGTTATCAGTGCAGTGGTAATGATGAACACACTTCTTATGAGAAACA 240
Qy 262 GGAGTCCGAGTTCTCTCTTTTGGATGTAACCTTCTCTCTGCTCCCAATATGGAACATGTA 321
Db 241 GGAGTCCGAGTTCTCTCTTTTGGATGTAACCTTCTCTCTGCTCCCAATATGGAACATGTA 300
Qy 322 CTAGCAGTTGCCAATGAAGAAGGCTTTGTTGATGTATATACAGAAATCAAAAGTTTC 381
Db 301 CTAGCAGTTGCCAATGAAGAAGGCTTTGTTGATGTATATACAGAAATCAAAAGTTTC 360
Qy 382 AGAAGAAGTCTCTCAAGAATGATGGCTCACTGGAATGCGTCTTTGACTTGGCCCTGG 441
Db 361 AGAAGAAGTCTCTCAAGAATGATGGCTCACTGGAATGCGTCTTTGACTTGGCCCTGG 420
Qy 442 GTTCTGCTGAACTTAAACTTTTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGAC 501
Db 421 GTTCTGCTGAACTTAAACTTTTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGAC 480
Qy 502 GTAAAGCTGTGAGCTGATTTGGAACATGCAAGGTCAATCAATGCAGCTCAAGTCAGTT 561
Db 481 GTAAAGCTGTGAGCTGATTTGGAACATGCAAGGTCAATCAATGCAGCTCAAGTCAGTT 540
Qy 562 GCCTTTTCTAAGTTTGAAAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG 621
Db 541 GCCTTTTCTAAGTTTGAAAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG 600
Qy 622 GTCTGGGATACAGGTGCAACAAAAAGATGGTTTTTATAGGCAAGTGAATCAAAATCAGT 681
Db 601 GTCTGGGATACAGGTGCAACAAAAAGATGGTTTTTATAGGCAAGTGAATCAAAATCAGT 660
Qy 682 GGAGCTCAAAATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATCA 741
Db 661 GGAGCTCAAAATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATCA 720
Qy 742 AAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAAAGTGTACTGTGGTCTCTTTTCAAGAC 801
Db 721 AAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAAAGTGTACTGTGGTCTCTTTTCAAGAC 780
Qy 802 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 861
Db 781 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 840
Qy 862 AAGATTTACTGCTTATTCGACNAGAAACCCATAGCATCCAGTCTTTCTGTACCCAGGT 921
Db 841 AAGATTTACTGCTTATTCGACNAGAAACCCATAGCATCCAGTCTTTCTGTACCCAGGT 900
Qy 922 AGCAGCACTCGAAAACTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981

901 AGCAGCACTCGAAAACTTGGATATTCAAGTCTGATTTTGGATTCCTCACTGGCTCTACTTTTA 960
982 TTTTGCTAAATTGCACAGCATACACTACATCTGCTTTAAATATGATGCTGGGTGAAGACTTCT 1041
961 TTTTGCTAAATTGCACAGCATACACTACATCTGCTTTAAATATGATGCTGGGTGAAGACTTCT 1020
1042 CCAGTGGCTATTATTTCAATGGACACACAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1101
1021 CCAGTGGCTATTATTTCAATGGACACACAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1080
1102 CCAGATGACCAAGTTTTAGTCACTGCTCAAGTGATGAAGCTGCCCTACATATGGAAGGTC 1161
1081 CCAGATGACCAAGTTTTAGTCACTGCTCAAGTGATGAAGCTGCCCTACATATGGAAGGTC 1140
1162 TCCACACCTTGGCAACCTCTACTGCTGCTCTGGGTCAATCTCAAGAGGTCAGCTCTGTG 1221
1141 TCCACACCTTGGCAACCTCTACTGCTGCTCTGGGTCAATCTCAAGAGGTCAGCTCTGTG 1200
1222 TGTGCTGTCTCCATCTGACTTTCACAAAGATTGCTACCTGTTCTGATGACAAATACACTAAAA 1281
1201 TGTGCTGTCTCCATCTGACTTTCACAAAGATTGCTACCTGTTCTGATGACAAATACACTAAAA 1260
1282 ATCTGCGCTTGAATAGAGCTTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCACGGTG 1341
1261 ATCTGCGCTTGAATAGAGCTTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCACGGTG 1320
1342 GGTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGCTAGTACAGTAAGAGTAGC 1401
1321 GGTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGCTAGTAAACAGTAAGAGTAGC 1380
1402 CAGAGTACTCTCCCAAGGCCCCAGGGTAAAGTGCAATCCATCCAAATCTTCCCGGTCA 1461
1381 CAGAGTACTCTCCCAAGGCCCCAGGGTAAAGTGCAATCCATCCAAATCTTCCCGGTCA 1440
1462 TCCGAGCTTGTGCCCCAAGCTGTGTGGAGACTCCCTCTTCCCTTCAAATATCTCTACG 1521
1441 TCCGAGCTTGTGCCCCAAGCTGTGTGGAGACTCCCTCTTCCCTTCAAATATCTCTACG 1500
1522 TTCTCTATTAAACCTCTCTGCAAGGCCCGGTCTCCCATGCAAGAGAGGCTCTGTG 1581
1501 TTCTCTATTAAACCTCTCTGCAAGGCCCGGTCTCCCATGCAAGAGAGGCTCTGTG 1560
1582 TCTCTCGTCTCTCCCAAGCCACTTCATCTTCAAGATGCGATTAGAACTGGGTGACC 1641
1561 TCTCTCGTCTCTCCCAAGCCACTTCATCTTCAAGATGCGATTAGAACTGGGTGACC 1620
1642 CGAACACCTTCTCTATCACCACCCATCACTCCACCTGCTTGGAGACCAAGATCATGTCT 1701
1621 CGAACACCTTCTCTATCACCACCCATCACTCCACCTGCTTGGAGACCAAGATCATGTCT 1680
1702 CGAGAAAAGCCCTTATCTGTGAGCGAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1761
1681 CGAGAAAAGCCCTTATCTGTGAGCGAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1740
1762 TCTAGAAAATAGATTAAGAGGCGCTAGACTCAAGCTGTCTGGAGAGTGTGAACAACAAAAG 1821
1741 TCTAGAAAATAGATTAAGAGGCGCTAGACTCAAGCTGTCTGGAGAGTGTGAACAACAAAAG 1800
1822 TGTGTGAAGAGTTGTAATCTGTGTGACTGTGAGCTTGTGAGGCAAGTGTGAATAATCTTCAATTG 1881
1801 TGTGTGAAGAGTTGTAATCTGTGTGACTGTGAGCTTGTGAGGCAAGTGTGAATAATCTTCAATTG 1860
1882 GATCTGTGCTCTGCTGTGTAACCAAGAGACCTTAGTAAGACCTCTAGAGTCTCTAGAGTCTTACC 1941
1861 GATCTGTGCTCTGCTGTGTAACCAAGAGACCTTAGTAAGACCTCTAGAGTCTCTAGAGTCTTACC 1920
1942 AATCAAGCAAAATTCAGAGAGCTGTACAGTATCTCAGAGCTCCGCTCCCTATCAGT 2001
1921 AATCAAGCAAAATTCAGAGAGCTGTGTACAGTATCTCAGAGCTCCGCTCCCTATCAGT 1980
2002 CCGTATGCTTCAGAAAGCTGTGGAACGCTACCTCTTCTTTGAGACCTTGTGGAGAGGG 2061
1981 CCGTATGCTTCAGAAAGCTGTGGAACGCTACCTCTTCTTTGAGACCTTGTGGAGAGGG 2040

2062 TCTGAAATCGTAGGCAAGAGAGAAATAGTTCCCCAGAGAAATAAAACTGGTTGTTGGCCATG 2121
2041 TCTGAAATCGTAGGCAAGAGAGAAATAGTTCCCCAGAGAAATAAAACTGGTTGTTGGCCATG 2100
2122 GCAGCCAAACGGAAGGCTGAGAATCCATCTCCACGAAGTCCGTCAATCCAGACACCCCAAT 2181
2101 GCAGCCAAACGGAAGGCTGAGAATCCATCTCCACGAAGTCCGTCAATCCAGACACCCCAAT 2160
2182 TCAGAGACACAGCGGAAAGACATTCGCAAGCCCGGTCCACCATCAGGCCAGCTCCATG 2241
2161 TCAGAGACACAGCGGAAAGACATTCGCAAGCCCGGTCCACCATCAGGCCAGCTCCATG 2220
2242 AGGAAAATCTGCACATACCTTCCATAGAAAGTCCCAAGAGGACTTCTGTGGTCTCTGAACAC 2301
2221 AGGAAAATCTGCACATACCTTCCATAGAAAGTCCCAAGAGGACTTCTGTGGTCTCTGAACAC 2280
2302 TCACAGAAATATAGATCTTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACAAAG 2361
2281 TCACAGAAATATAGATCTTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACAAAG 2340
2362 CTGAGCTTTGGTCCACTAAACAAAGATGAAAATACAAAGAGTACTCTATAACTCTGGTC 2421
2341 CTGAGCTTTGGTCCACTAAACAAAGATGAAAATACAAAGAGTACTCTATAACTCTGGTC 2400
2422 TTTTAAAGAAAGCTGCCCTTTTCAATTTTGTAGACAAAAATCTTTTCAACGCTGAAATGTACTAA 2481
2401 TTTTAAAGAAAGCTGCCCTTTTCAATTTTGTAGACAAAAATCTTTTCAACGCTGAAATGTACTAA 2460
2482 TCTGTGTTCTATACCATTAATGTATATGCAAGCTTCCCGAGGATGAATGTGTGTTTAAAT 2541
2461 TCTGTGTTCTATACCATTAATGTATATGCAAGCTTCCCGAGGATGAATGTGTGTTTAAAT 2520
2542 TCATAAGTAATTTGTGCTACTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTAAT 2601
2521 TCATAAGTAATTTGTGCTACTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTAAT 2580
2602 TCATCTCTCTATAATAATAGCATCCAGTTCATCGAGGCAAAAAACAAGTTTCTTGTTA 2661
2581 TCATCTCTCTATAATAATAGCATCCAGTTCATCGAGGCAAAAAACAAGTTTCTTGTTA 2640
2662 TCTGAAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
2641 TCTGAAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
2722 GGTGACTGAGAAATCTCTGCTGAAAGCCCTGGTTCTGTTCTGCTCCAAACATGTATA 2781
2701 GGTGACTGAGAAATCTCTGCTGAAAGCCCTGGTTCTGTTCTGCTCCAAACATGTATA 2760
2782 ATTTTATTTGAATACATAATCTTTTCACTATG 2814
2761 ATTTTATTTGAATACATAATCTTTTCACTATG 2793

RESULT 15

ADR25784
ID ADR25784 standard; DNA; 4221 BP.

XX ADR25784;

XX AC
XX XX
DT 21-OCT-2004 (first entry)

XX DE Breast cancer prognosis marker #1645.

XX ds; breast cancer; prognosis; gene expression; diagnosis.

XX OS Homo sapiens.

XX XX WO2004065545-A2.

XX XX 05-AUG-2004.

XX XX 15-JAN-2004; 2004WO-US001100.

PF

XX	15-JAN-2003; 2003US-00342887.	XX	622	GTCTGGGATACAGGTGCAACAAAAGATGCGTGTATATAGCAAGTGAATCAAAATCAGT	681
PR		XX	601		
XX	(ROSE-) ROSETTA INPHARMATICS LLC.	XX	682	GGAGCTCAAAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGAAAGAAACAGAAATTC	741
PA	(NECA-) NETHERLANDS CANCER INST.	XX	661		
XX		XX	742	AAAGGACTTGTCTCTCTGTGGATTTCAGCAAAAGTGTACTGTGGTCTCTTTCAAGAC	801
PI	Van't Veer LJ, He Y;	XX	721	AAAGGACTTGTCTCTCTGTGGATTTCAGCAAAAGTGTACTGTGGTCTCTTTCAAGAC	780
XX	WPI; 2004-593473/57.	XX	802	GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGAATAACAAAGTATGGGATTTAGCT	861
DR		XX	781	GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGAATAACAAAGTATGGGATTTAGCT	840
XX		XX	862	AAGATATTACTGCTTATCGCAAGAACCCATAGCATCCAAAGTCTTTCTCTGTACCAGGT	921
CC	The invention relates to a method of classifying a breast cancer patient	XX	841	AAGATATTACTGCTTATCGCAAGAACCCATAGCATCCAAAGTCTTTCTCTGTACCAGGT	900
CC	according to prognosis by determining the similarity between the level of	XX	922	AGCAGACTCGAAAACCTTGGATATTCAGTCTGATTTTGGATTTCCACTGGGCTCTACTTTA	981
CC	expression of each of five genes for which markers are listed in the	XX	901	AGCAGACTCGAAAACCTTGGATATTCAGTCTGATTTTGGATTTCCACTGGGCTCTACTTTA	960
CC	specification, in a cell sample taken from the breast cancer patient, to	XX	982	TTTGTAAATTCACAGACGATTAACATCTACATGTTTAAATATGACTGGGTGGAAGACTTCT	1041
CC	control levels of expression for each respective five genes to obtain a	XX	961	TTTGTAAATTCACAGACGATTAACATCTACATGTTTAAATATGACTGGGTGGAAGACTTCT	1020
CC	patient similarity value. The methods are useful for classifying a breast	XX	1042	CCAGTGGCTATTTTCAATGGACACACAGAACTCTACTTTTATGTAATAATCCAGCCTTAGT	1101
CC	cancer patient according to prognosis. Kits and computer program products	XX	1021	CCAGTGGCTATTTTCAATGGACACACAGAACTCTACTTTTATGTAATAATCCAGCCTTAGT	1080
CC	are useful for data analysis using the diagnostic, prognostic and	XX	1102	CCAGATGACCAAGTCTTGTAGTCAAGTGGCTCAAGTGAAGTGCCTACATATGGAAGTCT	1161
CC	statistical methods of the invention. This sequence corresponds to a	XX	1081	CCAGATGACCAAGTCTTGTAGTCAAGTGGCTCAAGTGAAGTGCCTACATATGGAAGTCT	1140
CC	marker used in the method of the invention.	XX	1162	TCCACACCTGGCACCTCTACTGTGCTCTGGGTCAATCTCAAGAGTCAAGTCTGTG	1221
XX	Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;	XX	1141	TCCACACCTGGCACCTCTACTGTGCTCTGGGTCAATCTCAAGAGTCAAGTCTGTG	1200
Query Match	98.5%; Score 2789.8; DB 13; Length 4221;	XX	1222	TGCTGGTCTCATCTGACTTTCACAAAGATTGCTACTCTCTGATGACAAATACACTAAAA	1281
Best Local Similarity	99.9%; Pred. No. 0;	XX	1201	TGCTGGTCTCATCTGACTTTCACAAAGATTGCTACTCTCTGATGACAAATACACTAAAA	1260
Matches 2791; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	XX	1282	ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGTGATAAACTTTCCACGGTG	1341
QY	22 CGATAACGATTGTGTGTGAGAGGCGCAACGTGCGATTTCGCTGAACTTGGAGGCATT	81	1261	ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGTGATAAACTTTCCACGGTG	1320
DB	1 CGATAACGATTGTGTGTGAGAGGCGCAAGCTGCGATTTCGCTGAACTTGGAGGCATT	60	1342	GGTTGGGCTCTCAGAGAAAGAGAGTCAAGACCTGGCTAGTAAACAGTAAACAGTACG	1401
QY	82 TCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTGGGTG	141	1321	GGTTGGGCTCTCAGAGAAAGAGAGTCAAGACCTGGCTAGTAAACAGTAAACAGTACG	1380
DB	61 TCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTGGGTG	120	1402	CAGAGTACTCTCTGCAAGGCTTAAAGTGAATTCATCCAAATCTTTCCCGCTCA	1461
QY	142 CTCGCGACGCCAGCTTGGCTGCTGAGAAATGATGCTTCAATACCCCTCTTCAA	201	1381	CAGAGTACTCTCTGCAAGGCTTAAAGTGAATTCATCCAAATCTTTCCCGCTCA	1440
DB	121 CTCGCGACGCCAGCTTGGCTGCTGAGAAATGATGCTTCAATACCCCTCTTCAA	180	1462	TCCGAGCTTGTGCCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATCTCTAG	1521
QY	202 TCCCTTCTGACTGGTTATCAGTGCAGTGTGATGATGAACACACTTCTTATGGAGAAACA	261	1441	TCCGAGCTTGTGCCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATCTCTAG	1500
DB	181 TCCCTTCTGACTGGTTATCAGTGCAGTGTGATGATGAACACACTTCTTATGGAGAAACA	240	1522	TTCTCTATTAAACCTCTCTGCGAAGGCGCGTCTCCCATCAACAGAGAGGCTCTGTG	1581
QY	262 GGAGTCCCAAGTCTCTCTTTTGGATGATCCTTCTCTTCTGCTCCCAATATGGAACATGTA	321	1501	TTCTCTATTAAACCTCTCTGCGAAGGCGCGTCTCCCATCAACAGAGAGGCTCTGTG	1560
DB	241 GGAGTCCCAAGTCTCTCTTTTGGATGATCCTTCTCTTCTGCTCCCAATATGGAACATGTA	300	1582	TCCTCCGTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGCGATTAGAAATCTGGGTGACC	1641
QY	322 CTAGCAGTTGCCAATGAAGGCTTGTGCGATTGATGATGAACACACTTCTTATGGAGAAACA	381	1561	TCCTCCGTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGCGATTAGAAATCTGGGTGACC	1620
DB	301 CTAGCAGTTGCCAATGAAGGCTTGTGCGATTGATGATGAACACACTTCTTATGGAGAAACA	360	1642	CGAACCTTCTCATCACCCACCTCATCTCCACTGTCTGGAGACCAAGATCATGTCT	1701
QY	382 AGAAGAAAGTGTCTCAAGAAATGATGCTCACTGGAATGCCGTCTTTGACCTGGCGCTGG	441	1621	CGAACCTTCTCATCACCCACCTCATCTCCACTGTCTGGAGACCAAGATCATGTCT	1680
DB	361 AGAAGAAAGTGTCTCAAGAAATGATGCTCACTGGAATGCCGTCTTTGACCTGGCGCTGG	420	1702	CCGAGAAAAGCCCTTATTTCTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTGTAG	1761
QY	442 GTTCTCTGGTGAATTTAAATCTTTTACAGCAGAGGCTGATCAACAGCCAAATTTTGGGAC	501			
DB	421 GTTCTCTGGTGAATTTAAATCTTTTACAGCAGAGGCTGATCAACAGCCAAATTTTGGGAC	480			
QY	502 GTAAAGCTGTGAGCTGATGGAACATGCAAGGCTCATCAATGACGCTCAAGTCAGTT	561			
DB	481 GTAAAGCTGTGAGCTGATGGAACATGCAAGGCTCATCAATGACGCTCAAGTCAGTT	540			
QY	562 GCCTTTCTTAAGTTTGAAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	621			
DB	541 GCCTTTCTTAAGTTTGAAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	600			

Db 2761 ATTTTATTGAAATACATAATCTTTTCACTATG 2793

Search completed: November 9, 2006, 06:21:28
Job time : 1693 secs

```
1681 CCAGAAAAAGCCCTTATTCCTGTGAGCCAGAGTCATCCCAAGCAGAGGCTTCTCTGAG 1740
1762 TCTAGAAATAGATAAGAGGAGCTAGACTCAAGCTGTCTGAGAGTGTGAACAAAAG 1821
1741 TCTAGAAATAGATAAGAGGAGCTAGACTCAAGCTGTCTGAGAGTGTGAACAAAAG 1800
1822 TGTGTGAAGAGTTGTAACTGTGTGACTGTGAGCTTGTATGGCCAAAGTGTGAATCTTCAATTTG 1881
1801 TGTGTGAAGAGTTGTAACTGTGTGACTGTGAGCTTGTATGGCCAAAGTGTGAATCTTCAATTTG 1860
1882 GATCTGTGTCTTGTCTGTGTAAACAGGAAGACCTTAGTAAGGACTCTCTAGTCTCTTACC 1941
1861 GATCTGTGTCTTGTCTGTGTAAACAGGAAGACCTTAGTAAGGACTCTCTAGTCTCTTACC 1920
1942 AAATCAAGCAAAATTTGAAGAGCTGTGTACAGTATCTCAGAGCTCCGTCTCCTATCAGT 2001
1921 AAATCAAGCAAAATTTGAAGAGCTGTGTACAGTATCTCAGAGCTCCGTCTCCTATCAGT 1980
2002 CCGTATGTCTCAGAAAGCTGTGAAACGCTACCTCTCTCTTTGAGACCTTGTGGAAGAGG 2061
1981 CCGTATGTCTCAGAAAGCTGTGAAACGCTACCTCTCTCTTTGAGACCTTGTGGAAGAGG 2040
2062 TCTGAAATGCTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACTGGTTGTTGGCCATG 2121
2041 TCTGAAATGCTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACTGGTTGTTGGCCATG 2100
2122 GCAGCCAAACGGAAGCTGAGAAATCCATCTCCAGGAAGTCCGTCTCCAGACACCCCAAT 2181
2101 GCAGCCAAACGGAAGCTGAGAAATCCATCTCCAGGAAGTCCGTCTCCAGACACCCCAAT 2160
2182 TCCAGAGACAGAGCGGAAGACATTTGCCAAGCCGTCACCATCACGCCAGCTCCATG 2241
2161 TCCAGAGACAGAGCGGAAGACATTTGCCAAGCCGTCACCATCACGCCAGCTCCATG 2220
2242 AGGAAATCTGCACATCTTCCATAGAAAGTCCAGGAGGACTTCTGTGTCTCTGAACAC 2301
2221 AGGAAATCTGCACATCTTCCATAGAAAGTCCAGGAGGACTTCTGTGTCTCTGAACAC 2280
2302 TCACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACAAAG 2361
2281 TCACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACAAAG 2340
2362 CTGAGCTTTGGTCCACTAAACAAAGATGAAATACAGAGTGACTCTATACTCTGCTC 2421
2341 CTGAGCTTTGGTCCACTAAACAAAGATGAAATACAGAGTGACTCTATACTCTGCTC 2400
2422 TTTAAGAAAGCTGCTTTTCAATTTTGTAGACAAATCTTTTCAACGCTGAAATGTACTAA 2481
2401 TTTAAGAAAGCTGCTTTTCAATTTTGTAGACAAATCTTTTCAACGCTGAAATGTACTAA 2460
2482 TCTGGTTCTACTACCAATATATATGACAGCTTCCGAGATGAATGCTGTGTTAAAT 2541
2461 TCTGGTTCTACTACCAATATATATGACAGCTTCCGAGATGAATGCTGTGTTAAAT 2520
2542 TCATAAGTAAATTTCTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2601
2521 TCATAAGTAAATTTCTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2580
2602 TCATCTTCTCTATAATAATGACATCCAGTTTCATGGAGGCAAAAAACAAGTTTCTGTGA 2661
2581 TCATCTTCTCTATAATAATGACATCCAGTTTCATGGAGGCAAAAAACAAGTTTCTGTGA 2640
2662 TCCGAAACCTTTCTATGCTCAGTGAAGATATCTGCCAGCCAGCATGAGGCTGTGAA 2721
2641 TCCGAAACCTTTCTATGCTCAGTGAAGATATCTGCCAGCCAGCATGAGGCTGTGAA 2700
2722 GGTGACTGAGAAATCCTCTGCTGAGAGCCCTGGTTCTGTTCTGCTCCACATGTATA 2781
2701 GGTGACTGAGAAATCCTCTGCTGAGAGCCCTGGTTCTGTTCTGCTCCACATGTATA 2760
2782 ATTTTATTGAAATACATAATCTTTTCACTATG 2814
```

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:08:54 ; Search time 46 Seconds
(without alignments)
1526.918 Million cell updates/sec

Title: US-10-726-160-2

Perfect score: 3877

Sequence: 1 MLFNSVLRQPLQGLVLRGWS.....TYFHRKSQDFCGPEHSTEL 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 80.*

1: Piri.*

2: Piri.*

3: Piri.*

4: Piri.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793.5	20.5	758	S51748	lethal(2)denticlel
2	370.5	9.6	727	T24284	hypothetical prote
3	340	8.0	810	T48835	lethal(2)denticlel
4	295.5	7.6	1526	AC2239	WD-40 repeat prote
5	286.5	7.4	490	T37884	transcription fact
6	279.5	7.2	934	AG1089	WD-40 repeat prote
7	271	7.0	676	AH2195	hypothetical prote
8	268	6.9	1189	AH2154	WD-repeat protei
9	261.5	6.7	1227	AC1810	WD-40 repeat prote
10	260	6.7	1747	AC1842	WD-40 repeat prote
11	259.5	6.7	1356	T18521	beta transducin-li
12	258.5	6.7	1258	AI2155	WD-repeat protei
13	258.5	6.7	1708	AI2155	WD-40 repeat prote
14	256.5	6.6	343	C84870	probable splicing
15	256	6.6	515	S19487	hypothetical prote
16	254.5	6.6	1683	AF2071	WD-40 repeat prote
17	252	6.5	1189	AI2493	WD-repeat protei
18	250	6.4	1711	AD1842	WD-40 repeat prote
19	248.5	6.4	1551	AB2410	WD-repeat protei
20	247.5	6.4	303	S45461	hypothetical prote
21	236.5	6.1	559	AB2202	hypothetical prote
22	235.5	6.1	1049	T42045	beta transducin-li
23	234	6.0	317	T46032	WD-40 repeat regul
24	233.5	6.0	606	T13152	WDRI protein - hum
25	232.5	6.0	342	AE2490	WD-repeat protei
26	230.5	5.9	677	AE1861	serine/threonine k
27	227.5	5.9	340	T39362	u5 snRNP-like prot
28	225.5	5.8	473	T33805	hypothetical prote
29	225.5	5.8	786	AG2375	WD-40 repeat-prote

RESULT 1

S51748

lethal(2)denticleless - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 05-Oct-2004

C;Accession: S51748

R;Kurzik-Dumke, U.; Neubauer, M.; Debes, A.

submitted to the EMBL Data Library, December 1994

A;Description: Identification of the novel Drosophila melanogaster heat shock gene, leth

A;Reference number: S51748

A;Accession: S51748

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-758 <KUR>

A;Cross-references: UNIPROT:Q24371; UNIPARC:UPI000012E18D; EMBL:X83414; NID:G603538; PID

C;Genetics:

A;Gene: FlyBase:1(2)dtl

A;Cross-references: FlyBase:FBgn0013548

A;Introns: 3/2

F;206-239/Domain: WD repeat homology <WD1>

F;349-383/Domain: WD repeat homology <WD2>

Query Match 20.5%; Score 793.5; DB 2; Length 758;

Best Local Similarity 29.2%; Pred. No. 1.7e+40;

Matches 224; Conservative 112; Mismatches 268; Indels 163; Gaps 25;

QY 49 PVPP-FGCTFSSAPNMEHVLAVANEEGFVRLYNT-----ESQSF-RKKCFKEWMAHWN 99

Db 38 PEPPISAKFANCDGYRHILAIANEDGKITLQDTQRNHQPERQSLVGPQC-----HFN 91

QY 100 AVFDLAWFGELKLVTAAGDQTAKFWDVWVAGELIG--TKGHQCSLSKVAFSFEKAVFC 157

Db 92 AVFDLEWAPQWRFSVSGSDHTARLMEVAGSGIRGLNSYVGHTRSVKSAFAFKRTDPAVFA 151

QY 158 TGRDGNIMVWTRCNKNDGFYQVQNOISGAHNTSDKQTPSKPKKONSKGLAPSDVDFQ 217

Db 152 TGRDGAIIWDIRANLNMDLTSRVNDCYSGHTGGPGIPVSORKORTRTPKAVAGTTSS 211

QY 218 SVTVLVFDENTILVSAVADGIIKWDLRKNYTAHQEPIASKSFLYPGSSSTRKLYSSL 277

Db 212 SITGLAFQNDTLISCGAGDGVIKWDLARNYTAHKELPLRHKLPYAGSSTFR-GFTNL 270

QY 278 ILDTSGSTLIFANCTDNIWNFMNTGLKSPVAIFNGHQNSTFFVKVSSLSGDDFLVSGSS 337

Db 271 IVDSAGTRLYANCMNTIYCNLASYSRPLACYKGLNLSFTFYIKSCLSPDGKYLISGSS 330

QY 338 DEAYTWKYSTWQPTVLLGHSEVTSVCWCPSTFKIATCSDDNTLKIWLNL----- 391

Db 331 DERAYIWNLDHAEPLVALAGHTVEVTCAWSSSHDPCIVTCSDDARHKIWRIGPDLDGL 390

QY 392 -----RGLBEKPGGDK----- 402

ALIGNMENTS

```
Db 391 SEAEAEKRGYGTASYVREPKAFPGSSGNHKNLRLDLESTPRSLKRLMDQNERTPGSVE 450
Qy 403 -----LSTVGWASQ-----KKESRRPGLVTVTSSQSTPAKAPRVK-CNPS 441
Db 451 KTTTKRSFLEMLGAVAGQETEAEPQOKRAKPLESRGRLFGPSSQSTACRHLQLQISNEE 510
Qy 442 NSSPSSAAACAPSCAGDLPPLPSNTPTSIKTSAPAKARSPINRRGSSVSPKPPSPFKMSI 501
Db 511 DASPSRKQKNSAEDV-----SPLHKLJSTP--SHSPLSE--NVNHNVTSPPTTS----- 557
Qy 502 RNVVTRTPSSSPITPAGSETKMSPRKALIPVQSQSOAE--ACSESRRVRKRLDSSC 559
Db 558 ----AAAAAADAALNPPPIISAALYSTSNL-PNYVLGDGAPHLGIMSPRKAKEKVD--W 611
Qy 560 LESVKQKCVKSCNVCVTELDQGVBNLHLDLCLAGNOEDLSKDSLGPTKSKIEGAGTSTS 619
Db 612 LTNIRQKLMGRHVTLSEKI-----SEEQADVLA---SPRLQSLRQSEC 655
Qy 620 EPPSPISP-----YASESCGTLPPLRPGCGSEMGKENSPPENKMWLLAAAKRAEN 674
Db 656 SPRIHASPRRIRISHTDGGGT-----PAGSSH--SHSQSQPKTP-----SSRNSET 702
Qy 675 P-----SPRSPSQPNRRQSGKTLPSVTI-----TPSSMR 707
Db 703 TLLRPFISQSSVPAEETTTNAAPSSSDPHPAVTAAPLPLSMR 749

RESULT 2
T24284
hypothetical protein T01C3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24284
R:Wild, A.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19869
A:Accession: T24284
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-727 <MIL>
A:Cross-references: UNIPROT:Q22059; UNIPARC:UPI0000081E3P; EMBL:Z78413; PIDN:CA01663.1;
A:Experimental source: clone T01C3
C:Genetics:
A:Gene: CBSP:T01C3.1
A:Map position: 5
A:Introns: 39/3; 88/1; 153/1; 306/1; 553/2; 654/2

Query Match 9.6%; Score 370.5; DB 2; Length 727;
Best Local Similarity 20.9%; Pred. No. 7.2e-15;
Matches 173; Conservative 127; Mismatches 279; Indels 249; Gaps 35;

Qy 22 QVP-----LQSLLTGQCNGNDHTSYGTGVPVPPFGCTSSAPNMEHVLAVANEEGF 75
Db 15 RYPSIYEELSKHYSPAGETDDHTW-----VTARFSPHLNQEHLVYMGDDPN 65
Qy 76 VRLYNTESQSFRCCKPKE-----WMAHNAVFDLAWVP-GELKLVTAAGDQTAQFWDVKA 129
Db 66 IGIFDVRKFDQDRSVPLEERQLYFFPAHDAIMDVGVGPQKESQIVSISGDSITRCWDLNQ 125
Qy 130 GELIGTCK--GHQCSLKSAVESKPEKAVCTGGRDGNIMVMDTRCNCKDGFVRQVNOIS 186
Db 126 STLDRKSQVFFGEGSVRSICFAPDPNPNVFTVGGDRDFQVKIMDMRVSTVTKMGEEDCRMAT 185
Qy 187 GAHTNSDKQTTPSK-----PKKQNSKGLAPSVDFOQSQSVTLVLFQDENTLVSAGA-VDGI 239
Db 186 ITYKTAHPK-PSKVLTSGLTPSKAKAKTIE-----GVKVTSLVFLDEHHVASASENADSG 239
Qy 240 IKWDLRKNYTAVRQSPIASKSFLYPGSSTRKLGYSLSILDSTGSLTGFANCTDDNIYMN 299
Db 240 IRVWDIRKPTRNGEOP--ARILKVTPTSNKSYGVTCLTILDRGNLRFASCTDSSIFEYS 297
Qy 300 MTGLKTSPP-----VAIFNGH 314
```

```
Db 298 VPSESVSPKFLVSHQLRKSQTNKSIITLWHSMSCASTTTLPTHVLDOYTKNLELSVNSYTGA 357
Qy 315 QNSTFYVKSLSPPDQFLVSGSSDEAAIYWKVSTPWQ-----PPTVL 356
Db 358 TIHNEYTVACSPVDVIACGSEDSERAVVMDLQDYNTMNDKRLPDDIDKRRTKLPRFSC 417
Qy 357 LHSQVETSVWCWCPDFTKIATCSDDNTLKTWRLNRLGLEEKPGDKLSTVGW-ASQKKKE 415
Db 418 DGHLLKQVLNVGW-SRGTTFWMS-DEGGVRIW-----SFRNRCTWKLDEDDT 464
Qy 416 SRPGLVTVTSSOSTPAKAPRVKCNPNSSPSAAACAPSCAGDLPPLPSNTPTPTFSIKTSPAK 475
Db 465 SYP-----TTSQELGLSEYKIKKFKILKESDEAMSCFD-----SISLSPQ 504
Qy 476 -----ARSPINRRGSSVSPKPPSPFKMSIRNWRTRTPSSSPPI-----TPPASETKI 524
Db 505 RADSSGLSGSPQKNRGSKRPI-----FESPLKSICTNSPK---PLRLNRSPRAMSKL 554
Qy 525 -MSPRKALIPYSQKSSQAEACSESRRVRKRLDSSLESVKQKCVKSCNVCTE--LDGQV 581
Db 555 SFSPPSPLOPTN-----SNNODLVGYRTP-----RQIRNKKKKNNPFFYNEHPTEGLP 601
Qy 582 ENLHLDLCLAGNQBDSLGLPTKSKIEGAGTISISEP----- 621
Db 602 NFVY-----DTFVKULGESSKSDIEDSGLSKLTGOKRIEDMWQTKGENVAITV 651
Qy 622 ----PSPISPVASESCGTLPPLRPGCGSEMGKENSPPENKMWLLAAAKRAENPSP 677
Db 652 RARLPS-VSEFGESAC-----SKVITEDER-----IALHSPKRL---VL 686
Qy 678 RSPSSQTP---NSRRQSGKTLPSVITP-----SSMRKICTYFHR 715
Db 687 KSSSTQSPSPNSK-----PKFIPMTPRKPMDKRPTSRNLLHYFKK 727

RESULT 3
T48835
lethal(2)identicleless related protein [imported] - Neurospora crassa
A:Alternate names: protein 68B2.190
C:Species: Neurospora crassa
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: T48835
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Faritmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24541
A:Accession: T48835
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-810 <SCH>
A:Cross-references: UNIPROT:Q9P6V7; UNIPARC:UPI000006B8C6; EMBL:ALJ35321; GSPDB:GN00112;
A:Experimental source: cosmid contig 68B2; strain 74
C:Genetics:
A:Gene: NCSP:68B2.190
A:Map position: 2
A:Introns: 236/1

Query Match 8.8%; Score 340; DB 2; Length 810;
Best Local Similarity 23.1%; Pred. No. 5.9e-13;
Matches 123; Conservative 73; Mismatches 149; Indels 188; Gaps 20;

Qy 49 PVPFGCTGFSSAPNMEHVLAVANESEGFVRLYNTESQSFRCCKCFKEMAHNAVFDLAWVP 108
Db 281 PTP-----FADHPKAKNLVAVGDEEGYVRLDQTQNEF--SQINMFSQAHGNAIIDKAFSD 334
Qy 109 GELKLVTAAGDQTAQFWDVVKAGELIGTCKGHOCSLKSVAF--SKFEKAVCTGGRDGNIM 166
Db 335 DLLLATASGDTGKVLDFVTVQOPIISILGHHTASLKQVRFQPGRGANSVLATTSRDGSVQ 394
Qy 167 VMDTFC-----NKKDGFYRQ-----VNOISGAHTNSDKQTPSPKPKKQNS 206
Db 395 IWDLCRGPGVDVAIRSEAGLHARLPRTNPGCVNSIYDAHS---RIPQ--KOARNL 448
Qy 207 KGLASVDF-----QQSVTVVLFDQ---ENTLVSAGAVDGIKWLKKNYAT 252
```

Db 449 LPCASNVDAIGEVPCRLGEVSVTALQPLPPCREHLFLTACADASVKLWDIRAVHTSR 508
Qy 253 RQPIASKSLFLYP---GSSTRKLGYSLLIDSTGSLFANCTDDNLYMFN-----299
Db 509 HHKYSTVPSFTAPLPSHANWRPFGISMALGGDSRLYTLCKDNTVYAYSTAHLVLGHAV 568
Qy 300 -MTGLKTSPIAIFNGHQN-----STFYVKGSLSP-----DDQFLV 333
Db 569 ELTAARPGEEPFRHNNHCTAHEGLGPLYGFRHPLFHATSFYVKAIRPAKNHGSSELLA 628
Qy 334 SGSSDEAAIYW-----KVSTPWQ-----PPT-----354
Db 629 VGSDDGAAVLFPDTERYIEKAFWQRHPNQEDDESVLVGEPTSCIPVTPARPLTRSAT 688
Qy 355 -----VLLGHSQEV 363
Db 689 TSLPLSSLSMSSSSGGGRPTSSASLFGIKGGHHSSNPIIRGTPLVRGHEKEV 748
Qy 364 TSVWCPSDFTKIATCSDNTLKIWRNLRGLEB---KPGGD---KLSTVGWA 409
Db 749 GALTWILDG--SLVTASDDYLVRWR--EGREEAADLRGTGETGRRWGCGWA 797
RESULT 4
AC2239
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AC2239
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC2239
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1526 <KUR>
A;Cross-references: UNIPROT:QBYR11; UNIPARC:UPI000013C28A; GB:BA000019; PIDN:BA075165.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr3466
Query Match 7.6%; Score 295.5; DB 2; Length 1526;
Best Local Similarity 24.5%; Pred. No. 6.5e-10;
Matches 107; Conservative 82; Mismatches 170; Indels 77; Gaps 17;
Qy 5 SVLRQPLGVLRNWSQYPLQSLITGYCQSGNDEHTSYGETGVPPFPGCTESSAPNME 64
Db 953 SVVFSNLSMLASGSD---QTVRLWDISSGELCIFQGH-----GWVYSVAFNLD 1001
Qy 65 -HVLAVANEFGFVRLYNTESSQSPKCKFEKMAHNAVFLAWPGLKLVTAAGDQTA 123
Db 1002 GSNLATGSGDQTVRLWDISS---SQCFYIFQGHSTCVRSVVFSSDGMALASGSDQTVR 1057
Qy 124 FWDVKAAGELIGTCGKHQCSLKSVAFKFAVCTGGRDGNIMVWDTNCKN----KDGFY 179
Db 1058 LWDISSGNCILYTLQGHSTCVRSVVFSP-DCAMLASGDDQIVRLWDISSGNCILYTLQGYT 1116
Qy 180 RQV-----NQISGAHNTSDK-----QTPSKPKKQNSKGLAPSVDFQSGVTVVLFDQE 227
Db 1117 SWRFLVFSNGVTLANGSDQIVRLWDISSK-----KCLYTLQGHNTWNVNAVAFSPD 1169
Qy 228 NTLVSAGAVDGIKWDLRKNYTAAYRQEPITAS--KGFY--PGSSTRKLGYSS-----275
Db 1170 GATLASGSGDQTVRLWDISSKCLYLQGHSTSWNSVVFNPDCSTLASGSDQTVRLWEI 1229
Qy 276 -----SLIDSTGSLFANCTDDNLYMFNMTGLKTSPIAIFNGHQNST 318
Db 1230 NSSKCLCTFQGHSTSWNSVVFNPDCGMLASGSDKTVRLWDISSKCL--LHTFQGHNTN-- 1285

Qy 319 FVYKS-SLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHSQEVTSVCWCPSPDFTKIA 377
Db 1286 -WVNSVAFNPDGSMALSGSDQTVRLWEISSKCLHT-FQGHSTWVSSVTFSP-DGTMLA 1342
Qy 378 TCSDDNTLKIWRNLRG 393
Db 1343 SGSDDDQTVRLWSISSG 1358
RESULT 5
T37884
transcription factor CDC10 target protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37884
R;Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z21751
A;Accession: T37884
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-490 <SKE>
A;Cross-references: UNIPROT:Q10990; UNIPARC:UPI000013AA75; EMBL:Z98597; PIDN:CAB11228.1;
A;Experimental source: strain 972h; cosmid ci7H9
C;Genetics:
A;Gene: cdt2; SPDB:SPAC17H9.19c
A;Map position: 1
Query Match 7.4%; Score 286.5; DB 2; Length 490;
Best Local Similarity 26.4%; Pred. No. 5.5e-10;
Matches 96; Conservative 67; Mismatches 144; Indels 57; Gaps 14;
Qy 52 PRGCTFSSAPNMEHVLAVANESGFVRLY-----NTESQSFRRKCKFEKMAHNAVND 103
Db 130 PFCLGFA---NNESLLAVCTETGALFLDSRFYDRQNEENQPSARR-IHGWLAHNAIFS 185
Qy 104 LAWVPGELKLVTAAGDQTAQFMDVKAGELI-----GTCKGHQCSLKSVAFSKFAVFC 157
Db 186 VNFSDKDSLLATSSGSDQTSKVPDLSTQOCITRLGREGVDGVHSHSVKQVNFCDNSPVNLV 245
Qy 158 TGRDGNIMVWDTNCRN--KKDG--FVRQVNIQSGAINTSDKQTPSPKPKKQNSKGLAPSV 213
Db 246 SCSRDGSIIFDMWRTHGITIDGHEHFQKPVLRIRKKAHENSGRDC-----288
Qy 214 DQQSVTVVLFDQENT--LVSAGAVDGIKWDLRKNYTAAYRQEPITASKSFYVPGSSTRK 271
Db 289 ----SITSATLWLPQSTQSVISSCSANSALKLWDLRTVHTV-RPLPAATTPEL--TTSKRD 341
Qy 272 LGVSSLLDSTGSLFANCTDDNLYMFNMTGLKTSPIAIFNGH--QNSTFYVKSLSLSPDD 329
Db 342 FGVNTVCTSPDGERIYAASRDSIIYEYSRHLNSGFCYTKDPLRLISSFYVKLACSPDG 401
Qy 330 QFLVSG-----SSDEAAIYWKVSTPWQPTVLL-GHSQEVTSVCWCPSPDFTKIATCSDNT 384
Db 402 ATLACGGVQDKTSGVVFDTTRNCSSSAMLGGHTKDVTAVDW--SSRQGLASISDDGS 459
Qy 385 LKIW 388
Db 460 VRWV 463
RESULT 6
AG1889
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AG1889
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AG1889
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-934 <KUR>
A:CROSS-references: UNIPROT:O8YZ23; UNIPARC:UPI00000CE10; GB:BA000019; PIDN:BA072622.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0664

Query Match 7.2%; Score 279.5; DB 2; Length 934;
Best Local Similarity 22.4%; Pred. No. 3.3e-09;
Matches 95; Conservative 67; Mismatches 170; Indels 93; Gaps 13;

QY 19 WSSQVPLQSLGTYQCSGNDEHSTYGETGVPPFGCTFSSAPNMEHVLAVANEGFVRL 78
DB 436 WREGKLLHTEGHDKVN-----SITFS--PDQG-LIATVGVNDTKL 476
QY 79 YNTESQSFKKCFKEMAHNAVFDLAWVPGLKLVTAAGDQTAKFWDVKAGELIGTCCKG 138
DB 477 WNLGKELR-----TFRGHQDMINSVSFSPDGKQIATASGDRTVKLWLSLDGKEL-OTLRG 530
QY 139 HQCSLKSVAFSKPEKAVFTCGGRDNIMVMDTRCNKKDGFYRQVNOISGAHNTSDKQTPS 198
DB 531 HQGVNSVTFSPDGKLI-ATASGDRTVKLWNSKGQELETLGHTD----- 574
QY 199 KFKKQNSKGLAPSDVDFQOSVTVVLFDENTLVSAGAVDGIIVKWDLRKNYTYAYRQEP 258
DB 575 -----AVNSVAFSPDGTSIATAGNDKTAIKWLN-----S 604
QY 259 SKSFLYPGSSRTRKLGYSLLDSTGLTFANCTDDNIYMFNMTGLKTSVPAIFNGHONST 318
DB 605 PMSIIVRHEDEVF---DLVFPNGKYIATASWDKTAUKWSIVGDKLQELRTFNGHQRV 661
QY 319 FYVKSLSLSDQFLVSGSSDEAAYIWKVSTPWPQPTVLLGHQSQVTSVCWCPSPDFTKIAT 378
DB 662 --NKLFSFDPGXVIATSTWDKAKLWNLDTLQ--KTLTGKDTVWSVNFSP-DGQLIAT 716
QY 379 CSDDNTLKIWLNRGLLE-----KPGDKLSTVGWASQKKESRPGLVTVT 424
DB 717 ASDEKTVKLWNRDGLLKTLPQSSVYNSAVFSPDGKLIATAGWDKTVKINSIDGRLOKT 776
QY 425 SSQST 429
DB 777 LTGHT 781

RESULT 7
AH2195
hypothetical protein alr3119 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2195
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:2159285; PMID:11759840
A:Accession: AH2195
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-676 <KUR>
A:CROSS-references: UNIPROT:O8YS66; UNIPARC:UPI00000CE698; GB:BA000019; PIDN:BA074818.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3119

Query Match 7.0%; Score 271; DB 2; Length 676;
Best Local Similarity 26.8%; Pred. No. 7.1e-09;
Matches 80; Conservative 48; Mismatches 115; Indels 56; Gaps 9;

QY 97 HNAVFDLAWPCELKLVTAAGDQTAKFWDVKAGELIGTCCKHQSLSKVAFSKPEKAVF 156

DB 390 HASDVNSVAFSPNGEFLASGDDKTIKVNWLNKQKIHTLPHSGWVWAIAPSPDGKTLA 449
QY 157 CTGGRDGNIMVMDTRCNKKDGFYRQVNOISGAHNTSDKQTSKPKKQNSKGLAPSDVQ 216
DB 450 STGA-DKTIKLNWLTGK-----EIRHLG----- 475
QY 217 QSVTVVLPQDENTLVSAGAVDGIIVKWDLRKNYTYAYRQEPISKSF--LYPGSSRTRKLG 274
DB 476 QGVASVAFSPDGKTLASGSLDKTIKLN-----PATGKEIRTLQEHSS-----GV 520
QY 275 SSLILDSTGLTFANCTDDNIYMFNMTGLKTSVPAIFNGHONSTFYVKSLSLSDQFLVS 334
DB 521 ANVAFSPDGKTLASGSLDKTIKLNWLTGK--IHTLKGH--SGLVMSVAFNSDSQTLAS 576
QY 335 GSSDEAAYIWKVSTPWPQPTVLLGHQSQVTSVCWCPSPDFTKIATCSDNTLKIWLNRG 393
DB 577 GSKDKTIKLNWLTGK-TIRTIRLGHSDKVNVAVPRDSTVLASGSDNTTIKLNWLTG 634

RESULT 8
AH2154
WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2154
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:2159285; PMID:11759840
A:Accession: AH2154
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1189 <KUR>
A:CROSS-references: UNIPROT:O8YTD1; UNIPARC:UPI00000CE56F; GB:BA000019; PIDN:BA074490.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2791

Query Match 6.9%; Score 268; DB 2; Length 1189;
Best Local Similarity 21.9%; Pred. No. 2.2e-08;
Matches 129; Conservative 94; Mismatches 233; Indels 132; Gaps 24;

QY 51 PFGCTFSSAPNMEHVLAVANEGFVRLYNTESQSFRRKCKPEKWA---HNAVFDLAWV 107
DB 703 PVAVVHFSFKGNM---IATASRDGTAKVWSLDG-----KELSLGCHKVWVYVNF 751
QY 108 PCELKLVTAAGDQTAKFWDVKAGELIGTCCKHQSLSKVAFSKPEKAVFTCGGRDGNIMV 167
DB 752 EDGKWLVTASDRDKTAIWDLQKEL-ATLRGHSDTVASAVFSR-DGQTIATASSDKTVRL 809
QY 168 WDTRCNKKDGFYRQVNOISGAHNTSDKQTSKPKKQNSKGLAPSDVDFQOSVTVVLPQDE 227
DB 810 WNRKGEELQVFWGHTDVAWG-----VNLSDKG 836
QY 228 NTLVSAGAVDGIIVKWDLRKNYTYAYRQEPISKSFYVKSLSLSDQFLVSDEAAIY 285
DB 837 KLLVSSGE-DGTVRLWNNGEAGKFQ-----SLSFNLGEAAAGTISFSPDGKILGTG 891
QY 286 LFANCTDDNIYMFNMTGLKTSVPAIFNGHONSTFYVKS-SLSPDQFLVSGSSDEAAIY 344
DB 892 TWAK-----LWNHQQ---QELVTLNGHSDT---LRLQFSFDQIATASRDKTVKLM 938
QY 345 KVSTPWPQPTVLLGHQSQVTSVCWCPSPDFTKIATCSDNTLKIWLNR----- 391
DB 939 NLN--GKERATLHGQADVRSATFSP-DSKTIASASWDTTVKLNLRGRIEIMTLRHOAG 995
QY 392 RGLKEKPGDKLSTVGWASQKKESRPGLVTVTSQSSTPAKAPVKCNPSNSSFSSAAC 450
DB 996 VRNVFSFPDQIIATASDGTAKLWNRQDELVT-----LKGHOAGIQAVSFSPSQVI 1049
QY 451 APSCAGDLPLPSNTPTFSIKTSPAK-ARSPINRRGVSVSVPKP-----PSSFKMSIRN 503

Db 1050 ATA-----SKDKTVKLNROGKELLTLTGHGEVNAVVSFSPNRETATASEDMTVKL 1101
QY 504 WYTRTSSSPPIPPASETKIMSPKALIPVSKSSQAEACSESRRNVRKRLDSSCLESV 563
Db 1102 WNLK-----GQGMQTLGLDAGVKSVSFSPDGKVLASDSGLGKVTLWNLDPDSS 1150
QY 564 KQKCV-KSCNCVTE--LDGQVENLHLDLCLAGNQEDLSKDSLGPRTKS 608
Db 1151 PEKLLAQCACVRYLLNSADIKHEHRLC-----DKLGTVKA 1188

RESULT 9
AE1810
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AE1810
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE1810
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1227 <KUR>
A:Cross-references: UNIPROT:Q8Z0R1; UNIPARC:UPI00000CDBE5; GB:BA000019; PIDN:BAB77553.1
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr0029

Query Match 6.7%; Score 261.5; DB 2; Length 1227;
Best Local Similarity 23.7%; Pred. No. 5.7e-08;
Matches 102; Conservative 65; Mismatches 166; Indels 97; Gaps 16;
QY 17 NGSSQVPLQSLTGVOCSDNDEHTSYGETGVPPPGCTFSAPNMEHVLAVANEGFV 76
Db 792 SGEDSTVRLMDVKTG-QCQVIFGHSHKKVYSVFSPDGQT-----LASCGBDRSI 840
QY 77 RLYNTESQSPRKKCFKEMMAHNAVFDLAWVPGLKLVTAAGDQTAKFWDVKAGELI--- 133
Db 841 KLWDIQ----RGECVNTLWGHSSQVWAIAPSPDRTLISCSDDQTAARLWDVITGNSLIL 896
QY 134 -----GTC-----KGHQSLSKSVAFSPEKAV 155
Db 897 RGYTRDVYSVAFSPDSQILASGRDDVTYIGLWNLKTGECHPLRGHQRIRSVAFHPDGK-I 955
QY 156 FCTGGRDGNIMWTRCNKKDGFVRQVNOISGAHN-----TSDKQTPSKPKK----- 202
Db 956 LASGSADNTIKLWDISDTHNSKIR---TLTGHNTVMWTVVFPDPKHTLASSSEDRIRL 1012
QY 203 --KONSKGLAPSVDFQOSVTVLWFQDENTLVLSAGAVDGIKVDLKNYTAAYRQEPATSK 260
Db 1013 WDXTDGLCKLKHSHWVTVAFSPDGRILASGSADSEIKIND-----VASG 1060
QY 261 SFLYPGSSTRKGLGYSSILSDSTGLFANCTDD-NIYFMNTGLKTSF-VAIFNGHONST 318
Db 1061 KCLQTLTDPQGMWIS--VAFSLDGLTLASASEQTVKLN--LKTGECVHTLKGHEKQV 1115
QY 319 FYVKSSLPDQFLVSGSSDEAAIYKVVSTPWPQPTVLLGHSEVTSVCWCPSPDFTKIAT 378
Db 1116 YSV--AFSPNGQIAAGSSEDTTVKLMWDISTGSCVDTLKHGHTAAIRSVAFSP-DGRLLAS 1172
QY 379 CSDNTLTKIW 388
Db 1173 GSEDEKIQWL 1182

RESULT 10
AC1842
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AC1842
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC1842
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1747 <KUR>
A:Cross-references: UNIPROT:Q8Z0Z0; UNIPARC:UPI00000CDDC1; GB:BA000019; PIDN:BAB77807.1
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0283

Query Match 6.7%; Score 260; DB 2; Length 1747;
Best Local Similarity 27.6%; Pred. No. 1.1e-07;
Matches 108; Conservative 64; Mismatches 157; Indels 62; Gaps 21;
QY 67 LAVANEEGVRLYNTESQSF-----RKCC---PKEMMAHNAVFDLAWVPGLKLVTAAGD 119
Db 1193 LATANADYTIKLYALDTSCLIVNNLQKCIQLIKTFPGHTDIVTVVVFSPDSKTIIVSSSLD 1252
QY 120 QTAKFWDVKAGELIGTCGHQSLSKSVAFSPEKAVFCTGGRDGNIMWTRCNKKDGFY 179
Db 1253 KTIKLRID-GSIINTWNAHGNVNSISFSPDGKMI-ASGGEDNLVKLWQA-----TNG-- 1304
QY 180 ROVNOISGAHN--TSDKQTP-----SKPKKQNSKG--LAPSVDFQOSVTVVLF- 224
Db 1305 HLKTLTGHERITSVKFSPDGKILASASGDKTIKFWTDGKFLKTIAAHNNQOVNIFS 1364
QY 225 QDENTLVLSAGAVDGIKVV-----DLRKNYTAAYRQEPATSKF-----LYPGSS-----T 269
Db 1365 SDSKTLVSAGA-DSTMKVWKIDTLIKTISG-RGEQIRDVTFSPDNKVVIASASSDKTVRI 1422
QY 270 RLKGY-----SSLLDSTGTLFANCTDNIYFMNTGLKTSFVAIFNGHONSTFYV 321
Db 1423 RQNLVYKOKSNVNSVSNFDPDGKTFASAGWDGNITWQRETLAHSLSLTQKQNIIITV 1482
QY 322 KESLSLDDQFLVSGSSDEAAIYKVVSTPWPQPTVLLGHSEVTSVCWCPSPDFTKIATCSD 381
Db 1483 --SYSDGKTIATASADNTIKLWDSQTO-QLKTLTGKHKDRITLTFHFDNQI-IASGSA 1538
QY 382 DNTLKIWRNLRG--LEEKPG-GDKLSTVGWA 409
Db 1539 DKTIKIWRVNDGQLLTLTGHNDVTSVNF 1569

RESULT 11
T18521
beta transducin-like protein - Podospora anserina
C:Species: Podospora anserina
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18521
R:Saupé, S.; Turcq, B.; Begueret, J.
Gene 162, 135-139, 1995
A:Title: A gene responsible for vegetative incompatibility in the fungus Podospora anserina
A:Reference number: Z18944; MUID:96009891; PMID:7557402
A:Accession: T18521
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1356 <SAU>
A:Cross-references: UNIPROT:Q00808; UNIPARC:UPI000012C621; EMBL:L28125; NID:9607002; PID
C:Genetics:
A:Gene: het-el
A:Introns: 761/3

Query Match 6.7%; Score 259.5; DB 2; Length 1356;
Best Local Similarity 25.1%; Pred. No. 8.6e-08;
Matches 94; Conservative 57; Mismatches 134; Indels 89; Gaps 14;

QY 35 SGNDHTSYGTGVPVPPFGCTFSSAPNMEHVLAVANBEGFVRLVNTESQSRKCKFKEM 94
DB 942 SGSDHT-----IKWDAASGT-----CTQTL 963
QY 95 MAHNAVFLAWVPGELKLVTAAGDQAKFMDVKGAGELIGTCKGHQCSLKSVAFSKFEKA 154
DB 964 EHGSSVLSVAFSPDQORVAGSGDKTIKIWDASGTCTQTLEGGGWSVAFSPDQOR 1023
QY 155 VFTCGGRDNIMWDFTRCNKKDGFYRQVNOIGAHNTSDKQTPSPKKQNSKGLAPSV- 213
DB 1024 V-ASGSDDKTIKIWDT---ASGTCTQTLEGGGWSVAFSPDQORVAGSGDDHTIKIW 1078
QY 214 -----DFQSVTVVLFDENTLVSAGAVDGIKWDLRKNYTAIRQEPFASKS 261
DB 1079 DAVSGTCTQTLEGGGWSVAFSPDQORVAGSGIDGTIKWDAASGTCTQTL----- 1124
QY 262 FLYPGSSTRKL-CYS----SLILDSTGTLFANCTDDNLYMFN-MTGLKTSFVAIFNGHO 315
DB 1125 ----GTCTQTLEGGGWSVAFSPDQORVAGSGIDGTIKWDAASGTCTQTL-----EGHG 1177
QY 316 NSTFYVKS-SLSPDDQFLVSGSSDEAAYIKWVSTPMQPTVLLGHSGEYTSVCMWCPDFT 374
DB 1178 G---WQSVAFSPDQORVAGSGSSDKTIKIWDASGTCTQT-LEGGGWSVAFSP-DGQ 1232
QY 375 KIATCSDDNTLKIW 388
DB 1233 RVASGSSDNTIKIW 1246
RESULT 12
AI2155
WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AI2155
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AI2155
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1258 <KUR>
A:Cross-references: UNIPROT:Q8YTC2; UNIPARC:UPI000013C036; GB:BA000019; PIDN:BA074499.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2800
Query Match 6.7%; Score 258.5; DB 2; Length 1258;
Best Local Similarity 22.6%; Pred. No. 9e-08;
Matches 93; Conservative 72; Mismatches 147; Indels 99; Gaps 16;
QY 67 LAVANEEGFVRLVNTESQSRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQAKFMD 126
DB 825 LASGSGDRTIKIWNHT---GELKTYIGHTNSVSYAISPDKILVSGSGDRTIKLWD 880
QY 127 VRAGELIGTCKGHQCSLKSVAFSKFEKAVFC----- 157
DB 881 CQTHCIKTLHGHTNEVCSVAFSPDQGTCLACVSLDQSVRLMNCRTGQCCLKAWYGNTDWAL 940
QY 158 -----TGGRDGNIMWDTNCKK- DGFYRQVNOIGAHNTSDKQTPSPKPKKQ 204
DB 941 PVAFSPDRQILASGNDKTVKLWDQTKYISLEGTDFYGIAPSPDQTLASASTDS 1000
QY 205 NSK-GLAPSVDPQ-----QSVTVVLFDENTLVSAGAVDGIKWDLR-----KNYTA 251
DB 1001 SVRLMNIWGQCQFILLEHTDWYAVVFPHPQGIKIATGSADCTVKLWNISTGQCLKLTSE 1060
QY 252 YRQ-----EPFASKSF-----LYPGSSTRKLG-----YSSLILDSTGSLT 286

DB 1061 HSDKILGMAWSPDQGLLASASADQSVRLWDCTGRCVGLIRGHSNRVYSATP--SPNGEI 1118
QY 287 FANC-TDNNIYMFNMTG---LKTSVAIFPNGHQNSTFFVVKSSLSDDQFLVSGSSDEAY 342
DB 1119 IATCSTDQTVKIMWQOQKCLKT-----LTGHTNWVFDI---AFSPDGKILASASHDQTVR 1171
QY 343 IWKVSTPMQPTVLLGHSGEYTSVCMWCPDFTKIATCSDDNTLKIWRNLNRG 393
DB 1172 IWDVNT-GKCHHICIGHLHLSVVAFSP-DGEVVASGSDQTVRIWNVKTG 1220
RESULT 13
AE1866
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AE1866
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE1866
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1708 <KUR>
A:Cross-references: UNIPROT:Q8Y2I2; UNIPARC:UPI00000CDD76; GB:BA000019; PIDN:BA072436.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0478
Query Match 6.7%; Score 258.5; DB 2; Length 1708;
Best Local Similarity 23.7%; Pred. No. 1.3e-07;
Matches 108; Conservative 61; Mismatches 168; Indels 119; Gaps 19;
QY 19 WSSQYPLQSLITGYOCSGNDHTSYGTGVPVPPFGCTFSSAPNMEHVLAVANBEGFVRL 78
DB 1215 WSREGKLLNTLSGH-----ND-----AVLGIAWTPDQGT-----LASVGADKNIKL 1255
QY 79 YNTESQSRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQAKFMDVKA----- 129
DB 1256 WNRDG-----KLLKTWQGHDDAILGVAMSPKGTIATASFDTIKLWNRQGNLLKTLSGH 1310
QY 130 -----GELTGTCKGHQCSLKSVAFSKFEKAVFCT 158
DB 1311 TAGVTAVTAVTSPNGETIGSASIDATILKWSPOGLLGLTLKGHSNWSVSP-DGRIFAS 1369
QY 159 GGRDGNIMWD-----TRCNKKDGFYRQVNOIGAHNTSDKQTPSKPKKQNSKGLAPSV 213
DB 1370 GSROKTVTLRWDEVLRLNPKGDG---NDWVTSISFSDSGETLAAASRDQTVKILSRHG 1425
QY 214 DFQ-----SVTVVLFDENTLVSAGAVDGIKWDLRKNYTAIRQEPFASKSFLYPG 266
DB 1426 KLLNTFKGTGTSIWCVAMSPNRCMIASAKDQTVKLW-----QDG 1466
QY 267 SSTRKL-GYSSLIID---STGTLFANCTDDNLYMFNMTG---LKTSPVAIFPNGHQNSTF 319
DB 1467 KILHTLQHQDAVLAVAWSSDQVTSAGKQKIVKWSQGGQLLHT-----LQGHDTAVN 1521
QY 320 YVKSSLSDDQFLVSGSSDEAAYIKWVSTPMQPTVLLGHSGEYTSVCMWCPDFTKIATC 379
DB 1522 WY--SFSPDGKILASVSDTTVICLW--SRDQGLLHTLKEHSRRVGVAWSP-DGQILASA 1576
QY 380 SDDNTLKIWRNLNRGLEEK-PG-GDKLSTVGVWASQKK 413
DB 1577 SIDGTVKLWNRDGSLSRNLPGDGSFISVSPDCK 1612
RESULT 14
C84870
probable splicing factor [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 05-Oct-2004
C;Accession: C84870
R;Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84870
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-343 <STO>
A;Cross-references: UNIPROT:O22826; UNIPARC:UPI00000AA2EC; GB:AB002093; NID:G2281093; PI
C;Genetics:
A;Gene: At2g43770
A;Map position: 2

Query Match 6.6%; Score 256.5; DB 2; Length 343;
Best Local Similarity 25.6%; Pred. No. 2.3e-08;
Matches 79; Conservative 44; Mismatches 117; Indels 69; Gaps 9;

QY 89 KCFKEMAHNAVFDLAWPGEIKLYTAAGDQAKFWDKAGELIGTCCKHQCSLKSVAF 148
DB 87 KNFMVLKGHKNALDLHMTSDGQIVSASPDKTVRAWDVETGKQIKQMAEHSFVNSCCP 146
QY 149 SKPEKAVFCTGGRDGNIMVWDTKCNKKGDFYRQVNOISGAHNTSDKQTPSKPKKQNSKG 208
DB 147 TRGGPPLIISGDDGTAKLWDMR-----QRC 172
QY 209 LAPSVDFPQQSVTVVLFQDENTLVSAAGVDGIIKVDLRKNYTAIROEPIASKSFLYPGSS 268
DB 173 AIQTFFPKYQITAVSFSDAADKIFTGGVNDVVKVMDLR-----GEA 214
QY 269 TRKL-GYSSLI-----LDSTGSTIFANCTDNIYFWNT--GLKTSVPAIFNGHNS--TF 319
DB 215 TMTLEGHQDTITGMSLSPDGSYLLTNGMDNKLVCWDMRPYAPQRCVKIPEGHQHFEN 274
QY 320 YVKSSLSPPDQFLVSSGSDAAIYIKVS---TPWQPPVLGHQSQVTSVCWCPSPDFTKI 376
DB 275 LLKCSWSPDGTKTAGSSDRMHWMTTSRTYIKLP----GHTGSVNECVFHPTE-PII 329
QY 377 ATCDDNTL 385
DB 330 GSCSSDKNI 338

RESULT 15
S19487
hypothenical protein YCR072c - Yeast (*Saccharomyces cerevisiae*)
C;Species: *Saccharomyces cerevisiae*
C;Date: 31-Mar-1992 #sequence_revision 04-Dec-1992 #text_change 05-Oct-2004
C;Accession: S19487; S26657
R;Ballesta, J.P.G.; Franco, L.; Hoenicka, J.; Jimenez, A.; Remacha, M.; Sanz, E.
submitted to the Protein Sequence Database, March 1992
A;Reference number: S19486
A;Accession: S19487
A;Molecule type: DNA
A;Residues: 1-484, 'I', 486, 'TKL', 490, 492-515 <BAL1>
A;Cross-references: UNIPROT:P25382; UNIPARC:UPI0000143E25; EMBL:X59720; MIPS:YCR072c
A;Note: this sequence has been revised in reference S26657
R;Ballesta, J.P.G.; Franco, L.; Hoenicka, J.; Jimenez, A.; Remacha, M.; Sanz, E.
submitted to the Protein Sequence Database, October 1992
A;Reference number: S26587
A;Accession: S26657
A;Molecule type: DNA
A;Residues: 481-503 <BAL2>
A;Cross-references: UNIPARC:UPI000017A4C5; EMBL:X59720; MIPS:YCR072c
C;Genetics:
A;Cross-references: SGD:S0000668
A;Map position: 3R
F;182-215/Domain: WD repeat homology <WD2>
F;225-265/Domain: WD repeat homology <WD3>
F;274-306/Domain: WD repeat homology <WD4>

F;398-431/Domain: WD repeat homology <WD5>
F;440-473/Domain: WD repeat homology <WD6>

Query Match 6.6%; Score 256; DB 2; Length 515;
Best Local Similarity 24.5%; Pred. No. 4.1e-08;
Matches 118; Conservative 64; Mismatches 192; Indels 108; Gaps 24;

QY 13 GVLK-NGWSSQVPLQSLTGYQCSGNDHTSYGETGVVP-PFGCTFFSSAPNMEHVLAV- 69
DB 47 GALLRPGAISEKQLEELN--QLNGTSDD-----PVPTFSCITQKKASDPVKITD 96
QY 70 -----ANERGFVRLYNTESSQSRKCKFKEMMA---HWNVAVFDLAWVP-GE 110
DB 97 ITDNLVSSLIKPGYNSTEDQITLLYTPRAVFKVPVTRSSSAIAGHSTILCSAFAPHTS 156
QY 111 LKLVTAAGDQAKFWDKAGELIGTCCKHQCSLKSVAFSKFAVCTGGRDGNIMVWDT 170
DB 157 SRMVTGAGDNTARIWDCDQTPMHTLKGHNWVLCVSWSP-DGEVIATGSMONTIRLWDP 215
QY 171 RCNKKDGFYRQVNOISGAHN---TSDKQTP-----SKPKKQNSKGLAPSV----- 213
DB 216 KSGQCLG-----DALRG-HSKWITSLSWEPIHLVKPGSKPRLASSSKDGTIKIWDIVSRV 269
QY 214 -----DFOQSVTVVLFQDENTLVSAAGVDGIIKVDLR-----KNYTAIROEPIA 258
DB 270 CQYTWSGHTNSVCVKWGGQGLYS--GSHDRTVRVMDINSQGRGINILKSHAHVWVHLSL 328
QY 259 SKSFLY-----PGSSTRKL--GYSSLLDSTGTLFANCTDNIYFWNT--TF 302
DB 329 STDYALRIGAFDHTGKKPSTPEEAQKALENTEKCKKNGNSEEMVVTASDDYTMFLMNP 388
QY 303 LK-TSPVAIFNGHNSSTFYVKSLSLPPDQFLVSSGSDAAIYIKW-----VSTPMQPPTV 355
DB 389 LKSTRPIARMTGKQLNVHV--AFSPDGRYIIVSASFDSIKLWDGDRGDKFEIST----- 439
QY 356 LIHGHSQEVTSVCWCPSPDFTKIATCSDDNTLKIW--RLNRLGLEKPG-GDKLSTVGWASQK 412
DB 440 FRGHIAVYQVAV--SSDCRLVSCSKDTTLKVDVTRKLSVDLPGHKDEVTVDWSDVG 498
QY 413 KK 414
DB 499 KR 500

Search completed: November 8, 2006, 18:14:41
Job time : 49 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:05:20 ; Search time 310 seconds
(without alignments)
2178.263 Million cell updates/sec

Title: US-10-726-160-2

Perfect score: 3877

Sequence: 1 MLFNSVLRLPQLGVLRLGWS.....TYFHRKSQDFCGPEHSTEL 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3877	100.0	730	Q9NZJ0_HUMAN	Q9nzj0 homo sapien
2	3872	99.9	730	Q9NM55_HUMAN	Q9nm55 homo sapien
3	3867	99.7	730	Q5V777_HUMAN	Q5v777 homo sapien
4	3499.5	90.3	729	Q3TLR7_MOUSE	Q3tlr7 mus musculus
5	3496.5	90.2	729	Q8BW38_MOUSE	Q8bw38 m 2 days pr
6	3479.5	89.7	729	Q80WY2_MOUSE	Q80wy2 mus musculus
7	3330.5	85.9	732	Q80WY1_MOUSE	Q80wy1 mus musculus
8	2835.5	73.1	597	Q3TTE9_MOUSE	Q3tte9 mus musculus
9	2705	69.8	534	Q6PAN1_MOUSE	Q6pan1 mus musculus
10	2293	59.1	434	Q9NM03_HUMAN	Q9nm03 homo sapien
11	2206.5	56.9	720	Q5ZJW8_CHICK	Q5zjw8 gallus gall
12	2098.5	54.1	713	Q6PLW0_XENTR	Q6plw0 xenopus tro
13	2052	52.9	710	Q4V837_XENLA	Q4v837 xenopus lae
14	2034.5	52.5	711	Q6GUP3_XENLA	Q6gup3 xenopus lae
15	1830.5	42.1	647	Q5RHI5_BRARE	Q5rhi5 brachydanio
16	1622.5	41.8	647	Q7ZU24_BRARE	Q7zu24 brachydanio
17	1615.5	41.7	647	Q8JH14_BRARE	Q8jh14 brachydanio
18	1255.5	32.4	475	Q4SCV3_TETNG	Q4scv3 tetraodon n
19	1216	31.4	232	Q9NM34_HUMAN	Q9nm34 homo sapien
20	1063	27.4	216	Q9C276_MOUSE	Q9c276 mus musculus
21	1020	26.3	211	Q96SN0_HUMAN	Q96sn0 homo sapien
22	793.5	20.5	769	L2DPTL_DROME	L2dptl drosophila
23	783	20.2	632	Q9PTB3_ANOGA	Q9ptb3 anopheles g
24	652	16.8	535	Q94C55_ARATH	Q94c55 arabidopsis
25	652	16.8	556	Q9LXK5_ARATH	Q9lxk5 arabidopsis
26	575.5	14.8	555	Q85L86_ORYSA	Q85l86 oryza sativ
27	404	10.4	676	Q61FK1_CAEBR	Q61fk1 caenorhabdi
28	370.5	9.6	727	Q22059_CAEEL	Q22059 caenorhabdi
29	340	8.8	810	Q9P6V7_NEUCR	Q9p6v7 neurospora
30	339	8.7	712	Q4WZRS_ASPFU	Q4wzrs aspergillus
31	332	8.6	727	Q2UDF0_ASPOR	Q2udf0 aspergillus

32	329	8.5	682	2	Q55LN8_CRYNE	Q55ln8 cryptococcus
33	324.5	8.4	717	2	Q5BBD2_EMENI	Q5bbd2 aspergillus
34	310.5	8.0	1661	2	Q3MC23_ANAVT	Q3mc23 anabaena va
35	305.5	7.9	1852	2	Q3MCV7_ANAVT	Q3mcv7 anabaena va
36	298.5	7.7	1789	2	Q3HJDI_TRIER	Q3hjd1 trichodesmi
37	295.5	7.6	1526	1	YY46_ANASP	Q8yrl1 anabaena sp
38	291	7.5	805	2	Q4PDT4_USTMA	Q4pdt4 ustilago ma
39	290	7.5	1211	2	Q47A03_DECAR	Q10990 schizosacch
40	286.5	7.4	490	1	CDT2_SCHPO	Q7nd05 gloeobacter
41	281.5	7.3	1193	2	Q7ND05_GLOVI	Q7nd05 gloeobacter
42	280.5	7.2	1173	2	Q4C9P2_CROWT	Q4c9p2 crocosphaer
43	279.5	7.2	934	2	Q8YZ23_ANASP	Q8yz23 anabaena sp
44	271	7.0	676	2	Q8YSG6_ANASP	Q8ysg6 anabaena sp
45	271	7.0	717	2	Q4I9G2_GIBZE	Q4i9g2 gibberella

ALIGNMENTS

RESULT 1

Q9NZJ0_HUMAN PRELIMINARY; PRT; 730 AA.

AC Q9NZJ0;

DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.

DT 01-OCT-2000, sequence version 1.

DT 07-FEB-2006, entry version 28.

DE L2DTL protein (RA-regulated nuclear matrix-associated protein)

DE (Retinoic acid-regulated nuclear matrix-associated protein)

DE (Denticleless homolog).

GN Names=L2DTL; Synonyms=DTL;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Mueller R., Ziegler B.L.;

RL Submitted (OCT-1999) to the EMBL/GenBank/DBDJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=21264461; PubMed=11278750; DOI=10.1074/jbc.M010802200;

RA Cheung W.M., Chu A.H., Chu P.W., Ip N.Y.;

RT "Cloning and expression of a novel nuclear matrix-associated protein

that is regulated during the retinoic acid-induced neuronal

differentiation.";

RL J. Biol. Chem. 276:17083-17091 (2001).

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Lymph, and Testis;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feilgenfeldt E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,

RA Diachenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalhus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [4]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Lymph, and Testis;
RG NIH MGC Project;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF195765; AAF35182.1; -; mRNA.
DR EMBL; AF345896; AAK54706.1; -; mRNA.
DR EMBL; BC033297; AAH33297.1; -; mRNA.
DR EMBL; BC033540; AAH33540.1; -; mRNA.
DR Ensembl; ENSG00000143476; Homo sapiens.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS00882; WD_REPEATS_2; 5.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 730 AA; 79469 MW; B4A149BC62059C4F CRC64;

Query Match 100.0%; Score 3877; DB 2; Length 730;
Best Local Similarity 100.0%; Pred. No. 3.3e-206;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFNSVLRQPQLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVVPVPPFGCTFSSA 60
DB 1 MLFNSVLRQPQLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVVPVPPFGCTFSSA 60
QY 61 PMEHVLA VANE EGFVRLYNTESQSFRKCKFKEMAHWNAVFDLAWVPGLKLVTAAGDQ 120
DB 61 PMEHVLA VANE EGFVRLYNTESQSFRKCKFKEMAHWNAVFDLAWVPGLKLVTAAGDQ 120
QY 121 TAKFWDV KAGELIGTCKGHQCSLKSVAFSKFEKAVFTCGRDNIMWDTRCNKKDGFYR 180
DB 121 TAKFWDV KAGELIGTCKGHQCSLKSVAFSKFEKAVFTCGRDNIMWDTRCNKKDGFYR 180
QY 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSQVTVVLFDENTLVLSAGAVDGI 240
DB 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSQVTVVLFDENTLVLSAGAVDGI 240
QY 241 KVDLRKNYTA VROPIAKSFYPGSSTRKLKGLYSSLLDSTGSLFANCTDDNIYMFNM 300
DB 241 KVDLRKNYTA VROPIAKSFYPGSSTRKLKGLYSSLLDSTGSLFANCTDDNIYMFNM 300
QY 301 TGLKTS PVAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB 301 TGLKTS PVAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
QY 361 QVTSVVCWCPSPDTKLTATCSDDNTLKIWRNLRGLEKPGDKLSTVGMASQKKESRPG 420
DB 361 QVTSVVCWCPSPDTKLTATCSDDNTLKIWRNLRGLEKPGDKLSTVGMASQKKESRPG 420
QY 421 VVTSSQSTPAKAPRVKCNPSNSSPSAAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
DB 421 VVTSSQSTPAKAPRVKCNPSNSSPSAAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
QY 481 NRRGVS SVSPKPPSPFKMSIRNWVTRTPSSSPITPPASETKIMSPKRALIPVSQKSSQ 540
DB 481 NRRGVS SVSPKPPSPFKMSIRNWVTRTPSSSPITPPASETKIMSPKRALIPVSQKSSQ 540
QY 541 AEACSESRNRVRRRLDSSCLLESYKQKVCVSCNCTVELDQGVENLHLDLCLAGNQEDLSK 600
DB 541 AEACSESRNRVRRRLDSSCLLESYKQKVCVSCNCTVELDQGVENLHLDLCLAGNQEDLSK 600
QY 601 DSLGPTKSKIEGAGTSISEPPSPISFYASESCGTLPLPLPCGEGSEVMGKENSPPENK 660
DB 601 DSLGPTKSKIEGAGTSISEPPSPISFYASESCGTLPLPLPCGEGSEVMGKENSPPENK 660
QY 661 NWLLAAAKRAKNAENPSRSPSSQTENSRQSGKTLPSPTITPSSMRKICTYFHRKSQBD 720
DB 661 NWLLAAAKRAKNAENPSRSPSSQTENSRQSGKTLPSPTITPSSMRKICTYFHRKSQBD 720

Db 661 NWLLAAAKRAKNAENPSRSPSSQTENSRQSGKTLPSPTITPSSMRKICTYFHRKSQBD 720
QY 721 FCGPEHSTEL 730
DB 721 FCGPEHSTEL 730

RESULT 2
Q9NWM5 HUMAN PRELIMINARY; PRT; 730 AA.
ID Q9NWM5; HUMAN
AC Q9NWM5;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Hypothetical protein FLJ20735.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK000742; BAA91355.1; -; mRNA.
DR Ensembl; ENSG00000143476; Homo sapiens.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS00882; WD_REPEATS_2; 5.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 730 AA; 79479 MW; 7640815C8D05303F CRC64;

Query Match 99.9%; Score 3872; DB 2; Length 730;
Best Local Similarity 99.9%; Pred. No. 6.2e-206;
Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLFNSVLRQPQLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVVPVPPFGCTFSSA 60
DB 1 MLFNSVLRQPQLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVVPVPPFGCTFSSA 60
QY 61 PMEHVLA VANE EGFVRLYNTESQSFRKCKFKEMAHWNAVFDLAWVPGLKLVTAAGDQ 120
DB 61 PMEHVLA VANE EGFVRLYNTESQSFRKCKFKEMAHWNAVFDLAWVPGLKLVTAAGDQ 120
QY 121 TAKFWDV KAGELIGTCKGHQCSLKSVAFSKFEKAVFTCGRDNIMWDTRCNKKDGFYR 180
DB 121 TAKFWDV KAGELIGTCKGHQCSLKSVAFSKFEKAVFTCGRDNIMWDTRCNKKDGFYR 180
QY 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSQVTVVLFDENTLVLSAGAVDGI 240
DB 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSQVTVVLFDENTLVLSAGAVDGI 240
QY 241 KVDLRKNYTA VROPIAKSFYPGSSTRKLKGLYSSLLDSTGSLFANCTDDNIYMFNM 300
DB 241 KVDLRKNYTA VROPIAKSFYPGSSTRKLKGLYSSLLDSTGSLFANCTDDNIYMFNM 300
QY 301 TGLKTS PVAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB 301 TGLKTS PVAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
QY 361 QVTSVVCWCPSPDTKLTATCSDDNTLKIWRNLRGLEKPGDKLSTVGMASQKKESRPG 420
DB 361 QVTSVVCWCPSPDTKLTATCSDDNTLKIWRNLRGLEKPGDKLSTVGMASQKKESRPG 420

Db 361 QEVTSVWCPSDFTKIATCDDNTLKIWLNRGLEBKPGDKLSTVGWASQKKESRPG 420

Qy 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAPSCAGDLPLPSNTPFTFSIKTSPAKARSPI 480

Db 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAPSCAGDLPLPSNTPFTFSIKTSPAKARSPI 480

Qy 481 NRRGVSVPKPPSPFQKSIIRNWVTRTPSSSPPIPPASETKIMSPRKALIPVSQKSSQ 540

Db 481 NRRGVSVPKPPSPFQKSIIRNWVTRTPSSSPPIPPASETKIMSPRKALIPVSQKSSQ 540

Qy 541 AEACSSNRVRRLDSSCLSVKQCVKSCNCTELDGOVENLHLDLCLAGNQBDLSK 600

Db 541 AEACSSNRVRRLDSSCLSVKQCVKSCNCTELDGOVENLHLDLCLAGNQBDLSK 600

Qy 601 DSGPTKSSKIEGAGTSISPPSPPIPIYASESGTLPPLRPGEGSEVMWGENSSPENK 660

Db 601 DSGPTKSSKIEGAGTSISPPSPPIPIYASESGTLPPLRPGEGSEVMWGENSSPENK 660

Qy 661 NWLLAMAAKKAENPSRPSSTQPNRRSQGKTLFSPVTITPSSMRKICTYFHRKSQED 720

Db 661 NWLLAMAAKKAENPSRPSSTQPNRRSQGKTLFSPVTITPSSMRKICTYFHRKSQED 720

Qy 721 FCGPEHSTEL 730

Db 721 FCGPEHSTEL 730

RESULT 3

Q5VT77 HUMAN PRELIMINARY; PRT; 730 AA.

AC Q5VT77;

DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.

DT 07-DEC-2004, sequence version 1.

DT 07-FEB-2006, entry version 13.

DE RA-regulated nuclear matrix-associated protein.

GN Name=RAMP; ORFNames=RP11-15111.1-001;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Almeida J.;

RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RA Sycamore N.;

RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

CC EMBL; AL606468; CAH70697.1; -; Genomic DNA.

DR EMBL; AC092814; CAH70697.1; JOINED; Genomic DNA.

DR EMBL; AL592297; CAH70697.1; JOINED; Genomic DNA.

DR EMBL; AL592297; CAH73803.1; -; Genomic DNA.

DR EMBL; AC092814; CAH73803.1; JOINED; Genomic DNA.

DR EMBL; AL606468; CAH73803.1; JOINED; Genomic DNA.

DR Ensembl; ENSG00000143476; Homo sapiens.

DR InterPro; IPR001680; WD40.

DR Pfam; PF00400; WD40; 5.

DR PRINTS; PR00320; GPROTEINBRPT.

DR ProDom; PD000018; WD40; 1.

DR SMART; SM00320; WD40; 5.

DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.

DR PROSITE; PS50082; WD_REPEATS_2; 5.

DR PROSITE; PS50294; WD_REPEATS_REGION; 1.

KW Repeat; WD repeat.

SQ SEQUENCE 730 AA; 79468 MW; CE8D54234D44F002 CRC64;

Query Match 99.7%; Score 3867; DB 2; Length 730;

Best Local Similarity 99.7%; Pred. No. 1.2e-205;

Matches 728; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MLFNSVLRQPLQGLVLRNGWSSQYPIQLSLTGTQCSGNDEHTSYGETGVVPVPGCTFSSA 60

Db 1 MLFNSVLRQPLQGLVLRNGWSSQYPIQLSLTGTQCSGNDEHTSYGETGVVPVPGCTFSSA 60

Qy 61 PNMEHLVAVANEEGFVRLYNTESQSFRKCKFKEMAHNAVFDLAWVPGLKLVTAAGDQ 120

Db 61 PNMEHLVAVANEEGFVRLYNTESQSFRKCKFKEMAHNAVFDLAWVPGLKLVTAAGDQ 120

Qy 121 TAKFMDVKAGELIGTCKGHQCSLSKVAFSKPKFAVFCCTGGRDGNIMWTDRCNKKDGFR 180

Db 121 TAKFMDVKAGELIGTCKGHQCSLSKVAFSKPKFAVFCCTGGRDGNIMWTDRCNKKDGFR 180

Qy 181 QYNQISGAHNTSDKOTPSKPKKQSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGI 240

Db 181 QYNQISGAHNTSDKOTPSKPKKQSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGI 240

Qy 241 KYWDLRKNTAYRQBPFIASKSFLYPCSSSTRKLGYSLLDSTGLFANCTDDNIYFMNM 300

Db 241 KYWDLRKNTAYRQBPFIASKSFLYPCSSSTRKLGYSLLDSTGLFANCTDDNIYFMNM 300

Qy 301 TGLKTSPIVAIFNGHONSFTYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360

Db 301 TGLKTSPIVAIFNGHONSFTYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360

Qy 361 QEVTSVWCPSDFTKIATCDDNTLKIWLNRGLEBKPGDKLSTVGWASQKKESRPG 420

Db 361 QEVTSVWCPSDFTKIATCDDNTLKIWLNRGLEBKPGDKLSTVGWASQKKESRPG 420

Qy 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAPSCAGDLPLPSNTPFTFSIKTSPAKARSPI 480

Db 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAPSCAGDLPLPSNTPFTFSIKTSPAKARSPI 480

Qy 481 NRRGVSVPKPPSPFQKSIIRNWVTRTPSSSPPIPPASETKIMSPRKALIPVSQKSSQ 540

Db 481 NRRGVSVPKPPSPFQKSIIRNWVTRTPSSSPPIPPASETKIMSPRKALIPVSQKSSQ 540

Qy 541 AEACSSNRVRRLDSSCLSVKQCVKSCNCTELDGOVENLHLDLCLAGNQBDLSK 600

Db 541 AEACSSNRVRRLDSSCLSVKQCVKSCNCTELDGOVENLHLDLCLAGNQBDLSK 600

Qy 601 DSGPTKSSKIEGAGTSISPPSPPIPIYASESGTLPPLRPGEGSEVMWGENSSPENK 660

Db 601 DSGPTKSSKIEGAGTSISPPSPPIPIYASESGTLPPLRPGEGSEVMWGENSSPENK 660

Qy 661 NWLLAMAAKKAENPSRPSSTQPNRRSQGKTLFSPVTITPSSMRKICTYFHRKSQED 720

Db 661 NWLLAMAAKKAENPSRPSSTQPNRRSQGKTLFSPVTITPSSMRKICTYFHRKSQED 720

Qy 721 FCGPEHSTEL 730

Db 721 FCGPEHSTEL 730

RESULT 4

Q3TLR7 MOUSE PRELIMINARY; PRT; 729 AA.

ID Q3TLR7_MOUSE

AC Q3TLR7;

DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.

DT 11-OCT-2005, sequence version 1.

DT 07-FEB-2006, entry version 6.

DE Mammary gland RCB-0526 Jyg-MC(A) cDNA, RIKEN full-length enriched

DE library, clone:G830048A21 product:L2DTL protein, full insert sequence.

GN Name=Dtl;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muroidae; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE.

CC TISSUE=Mammary gland;

RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Banai M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.P., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kello J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liut S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petkovsky N., Piazza S., Reid J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Sempole C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugtara K., Sultana R., Takenaka Y., Taki K.,
RA Tamoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nieuwen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zaborovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashina T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=23354663; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Mulcais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perle G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Vercato R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Harai A., Hashizume W., Imotani K., Iehli Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu N., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa K., Tanaka T., Matsura S., Kawai J.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1157-1171(2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Tida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBSJ databases.
CC -----

RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566 (2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oatono N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matuda H., Batalov S., Reisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher T.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmerland S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongave A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sadelin A., Schneider C., Semple C.A., Setou M., Shinada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX MEDLINE=20330913; PubMed=1076861; DOI=10.1101/gr.152600;
RA QVNOI SGAHNTSDKOTPSKPKKQNSKGLAPAVDSQQSVTVVLFQDENTLVSAGAVDGI 240
DB QVNOI SGAHNTADKOTPSKPKKQNSKGLAPAVDSQQSVTVVLFQDENTLVSAGAVDGI 240

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Iehikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [9]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC -----
DR EMBL; AK054412; BAC35769.1; -; mRNA.
DR EMBL; AK133177; BAE21543.1; -; mRNA.
DR Ensembl; ENSMUSG00000037474; Mus musculus.
DR MGI; MGI:1924093; Dcl.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00578; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS00882; WD_REPEATS_2; 5.
DR PROSITE; PS02994; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 729 AA; 79131 MW; C828FAFBA9929360 CRC64;
Query Match 90.2%; Score 3496.5; DB 2; Length 729;
Best Local Similarity 89.5%; Pred.No. 3.7e-185;
Matches 653; Conservative 36; Mismatches 40; Indels 1; Gaps 1;
QY 1 MLFNSVLQPQLGVLRNGWSSQYPLQSLITGQCSCNDEHTSYGETGVPPFGCTFSSA 60
DB 1 MLFNSVLQPQLGVLRNGWSSHYPLQSLISGQCNCNDEHTSYGETGVPPFGCTFCTA 60
QY 61 PMEHVLAVANEEGVRLYNTESQSFRRKCFKFWMAHNAVDFLAWVPGLKLVTAAGDQ 120
DB 61 PMEHVLAVANEEGVRLYNTESQSFRRKCFKFWMAHNAVDFLAWVPGLKLVTAAGDQ 120
QY 121 TAKFWDVWVAGELIGTCCKHQCSLKSVAFKPKVAFCTGGRDNTMVDTRCNCKDGFYR 180
DB 121 TAKFWDVWVAGELIGTCCKHQCSLKSVAFKPKVAFCTGGRDNTMVDTRCNCKDGFYR 180
QY 181 QVNOI SGAHNTSDKOTPSKPKKQNSKGLAPAVDSQQSVTVVLFQDENTLVSAGAVDGI 240
DB 181 QVNOI SGAHNTADKOTPSKPKKQNSKGLAPAVDSQQSVTVVLFQDENTLVSAGAVDGI 240

```
QY 241 KWDLRKNYTAHQEPIASKSFLYPGSSSTRKLGYSLLILDSTGSLTIFANCTDDNIYFMN 300
DB 241 KWDLRKNYTAHQEPIASKSFLYPGSSSTRKLGYSLLILDSTGSLTIFANCTDDNIYFMN 300
QY 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYKRWSTPWPQPTVLLGHS 360
DB 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYKRWSTPWPQPTVLLGHS 360
QY 361 QEVTSVCWCPSPDFTKIATCDDNTLKIWLNRGLBEKPGDGKLSLVGWSAQKKESRPGL 420
DB 361 QEVTSVCWCPSPDFTKIATCDDNTLKIWLNRGLBEKPGDGKLSLVGWSAQKKESRPGL 420
QY 421 VVTSSQSTPAKAPRVKCNPSNPSAACPACAGDPLPSPNTPTFSIKTSPAKARSP 480
DB 421 VVTSSQSTPAKAPRVKCNPSNPSAACPACAGDPLPSPNTPTFSIKTSPAKARSP 480
QY 481 NRRGVSVPKPPSPFMSIRNWTPTSSSPPIPPASSETKIMSPRKALIPVSQKSSQ 540
DB 481 NRRGVSVPKPPSPFMSIRNWTPTSSSPPIPPASSETKIMSPRKALIPVSQKSSQ 540
QY 541 AEACSESRNRVXRRLDSSCLSVKQKCVKSCNCTVLDGQVENLHLDLCLAGNQEDLSK 600
DB 541 AEACSESRNRVXRRLDSSCLSVKQKCVKSCNCTVLDGQVENLHLDLCLAGNQEDLSK 600
QY 601 DSI LGPTKSSKIEGAGTSISEPPSPISPYASESGTLPLPLRPGEGSEMVKGKENSSENK 660
DB 601 DSI LGPTKSSKIEGAGTSISEPPSPISPYASESGTLPLPLRPGEGSEMVKGKENSSENK 660
QY 661 NMLLAAMAARKAENSPRSPSSQTPNSRRQSGKTLSPVITPSSMRKICTYFHRKSQED 720
DB 661 NMLLAAMAARKAENSPRSPSSQTPNSRRQSGKTLSPVITPSSMRKICTYFHRKSQED 720
QY 721 FCGPEHSTEL 730
DB 721 FCGPEHSTEL 729

RESULT 6
Q80WY2 MOUSE
ID Q80WY2_MOUSE PRELIMINARY; PRT; 729 AA.
AC Q80WY2;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Retinoic-acid regulated nuclear matrix-associated protein.
GN Name=2810047L02Rik; Synonyms=Dtl, ramp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BALB/c;
RA Uenaka A., Hirano Y., Hata H., Sanda W., Aji T., Tanaka M., Ono T.,
RA Skipper J., Shimizu K., Nakayama E.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; AB095735; BAC76404.1; -, mRNA.
DR Ensembl; ENSMUSG0000037474; Mus musculus.
DR MGI; MGI:1924093; 2810047L02Rik.
DR MGI; MGI:1924093; Dtl.
DR InterPro; IPR001690; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
```

```
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 729 AA; 79161 MW; 81DAF4BACD5F38B1 CRC64;

Query Match      89.7%; Score 3479.5; DB 2; Length 729;
Best Local Similarity 89.2%; Pred. No. 3.2e-184;
Matches 651; Conservative 36; Mismatches 42; Indels 1; Gaps 1;

QY 1 MLFNSVLRLPOLGVLNRGWSOYPLQSLLTGYQCSGNDEHTSYGETGVVPVPPGCTFSSA 60
DB 1 MLFNSVLRLPOLGVLNRGWSOYPLQSLLTGYQCSGNDEHTSYGETGVVPVPPGCTFCTA 60
QY 61 PMEHLVAVANEEGFVRLYNTESQSFRKKCFKEMAHNAVFDLAWVPVGGELKLVTAAGDQ 120
DB 61 PMEHLVAVANEEGFVRLYNTESQSFRKKCFKEMAHNAVFDLAWVPVGGELKLVTAAGDQ 120
QY 121 TAKFWDVAVAGELIGTCKGHQCSLSKVAFSPKFAVCTCGRDGNIMWDTCKNKKDGFYR 180
DB 121 TAKFWDVAVAGELIGTCKGHQCSLSKVAFSPKFAVCTCGRDGNIMWDTCKNKKDGFYR 180
QY 181 QYNQISGAHNTSDKOTPSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGLI 240
DB 181 QYNQISGAHNTSDKOTPSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGLI 240
QY 241 KYWDLRKNYTAHQEPIASKSFLYPGSSSTRKLGYSLLILDSTGSLTIFANCTDDNIYFMN 300
DB 241 KYWDLRKNYTAHQEPIASKSFLYPGSSSTRKLGYSLLILDSTGSLTIFANCTDDNIYFMN 300
QY 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYKRWSTPWPQPTVLLGHS 360
DB 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYKRWSTPWPQPTVLLGHS 360
QY 361 QEVTSVCWCPSPDFTKIATCDDNTLKIWLNRGLBEKPGDGKLSLVGWSAQKKESRPGL 420
DB 361 QEVTSVCWCPSPDFTKIATCDDNTLKIWLNRGLBEKPGDGKLSLVGWSAQKKESRPGL 420
QY 421 VVTSSQSTPAKAPRVKCNPSNPSAACPACAGDPLPSPNTPTFSIKTSPAKARSP 480
DB 421 VVTSSQSTPAKAPRVKCNPSNPSAACPACAGDPLPSPNTPTFSIKTSPAKARSP 480
QY 481 NRRGVSVPKPPSPFMSIRNWTPTSSSPPIPPASSETKIMSPRKALIPVSQKSSQ 540
DB 481 NRRGVSVPKPPSPFMSIRNWTPTSSSPPIPPASSETKIMSPRKALIPVSQKSSQ 540
QY 541 AEACSESRNRVXRRLDSSCLSVKQKCVKSCNCTVLDGQVENLHLDLCLAGNQEDLSK 600
DB 541 AEACSESRNRVXRRLDSSCLSVKQKCVKSCNCTVLDGQVENLHLDLCLAGNQEDLSK 600
QY 601 DSI LGPTKSSKIEGAGTSISEPPSPISPYASESGTLPLPLRPGEGSEMVKGKENSSENK 660
DB 601 DSI LGPTKSSKIEGAGTSISEPPSPISPYASESGTLPLPLRPGEGSEMVKGKENSSENK 660
QY 661 NMLLAAMAARKAENSPRSPSSQTPNSRRQSGKTLSPVITPSSMRKICTYFHRKSQED 720
DB 661 NMLLAAMAARKAENSPRSPSSQTPNSRRQSGKTLSPVITPSSMRKICTYFHRKSQED 720
QY 721 FCGPEHSTEL 730
DB 721 FCGPEHSTEL 729

RESULT 7
Q80WY1 MOUSE
ID Q80WY1_MOUSE PRELIMINARY; PRT; 732 AA.
AC Q80WY1;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Meth A retinoic-acid regulated nuclear matrix-associated protein.
GN Name=2810047L02Rik; Synonyms=Dtl, Meth A ramp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
```



```
QY 553 RRLDSSCLSVKQKVCNCVTELDGQVENLHLDLCLAGNOEDLSKDSLGPTKSSKIE 612
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 420 RRLDSSCLSVKQKVCNCVTELDGQVAESLRLDCLCLSGTQEVLSQDSGPTKSSKTE 479
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 613 GAGTISSEPPSPISPVASESCGTLPLPLPCCGEGSEWGVKENSPPNKWLLAMAARKA 672
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 480 GAGTISSEPPSPVSPVASECGCLPLPLPCCGEGSEWGVKENSPPNKWLLAIAARKA 539
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 673 ENPSRPSQTPNSRRSQSGKTLPSVTTIPSSMRKICTYFHRKQEDFCGPEHSTEL 730
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 540 ENSRPSRPSQTPNSRRSQSGKTLPSVTTIPSSMRKICTYFHRKQEDFCGPEHSTEL 597
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
Q6PAN1_MOUSE PRELIMINARY; PRT; 594 AA.
AC O6PAN1;
CT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DE 07-FEB-2006, entry version 19.
DE 2810047L02Rik protein.
GN Name=Dtl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zerbini B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.E.,
RA Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Warra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC
CC EMBL, BC060208, AAH60208.1; -, mRNA.
DR Ensembl; ENSMUSG00000037474; Mus musculus.
DR MGI; MGI:1924093; Dtl.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 3.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD0000018; WD40; 1.
DR SMART; SM00320; WD40; 3.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE; PS50082; WD_REPEATS_2; 3.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 594 AA; 64233 MW; 180FD4B69B4A823 CRC64;
```

```
Query Match 69.8%; Score 2705; DB 2; Length 594;
Best Local Similarity 87.1%; Pred. No. 1.8e-14;
Matches 514; Conservative 28; Mismatches 42; Indels 6; Gaps 2;

QY 141 CSLKSAFSAKPEKAVFCGTRDGNIMWDTNCRNKKDGFYRVQVQISGAHNTSDKOTPSKP 200
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 CSLSQVS-----KSCVLYRGERGNIMWDTNCRNKKDGFYRVQVQISGAHNTADKOTPSKP 65
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 201 KKKQNSKGLAPSVDFQQSVTVVLFQDENTLVLSAGAVDGIWKWDLRKNKYAYRQBPISK 260
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 66 KKKQNSKGLAPAVDSQQSVTVVLFQDENTLVLSAGAVDGIWKWDLRKNKYAYRQBPISK 125
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 261 SFLYPGSSSTRKLGYSLLDSTGSTLFPANCTDDNIMFNMTGLKTSPIAFNGHONSTFY 320
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 SFLYPGTSSTRKLGYSLLDSTGSTLFPANCTDDNIMFNMTGLKTSPIAVFNGHONSTFY 185
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 321 VKSSLSRDDQFLVSGSSDEAAIYKVKSTPWPPTVLLGHSGQVTSVVCPCPSDFTKIATCS 380
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 VKSSLSRDDQFLVSGSSDEAAIYKVKSNPWPPTVLLGHSGQVTSVVCPCPSDFTKIATCS 245
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 381 DNTLKIWLNRGLLEKPGDKLSTVGWASQKKESRFGLVTVTSSQSTPAKAPVKCP 440
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 DNTLKIWLNRGLLEKEP-GDKHSIVGWTQKKKEVKACPVTPVSSQSTPAKAPRAKSP 304
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 441 SNSPSSAACAPSCAGDLPLPSNTPTFSIKTSPAKARSPINRGSVSVSPKPPSPFKMS 500
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 SISPSSAACTPSCAGDLPLPSSTPTFSVKTPTATTSVSRSGSISVSPKPLSPFKMS 364
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 501 TRNWTRTPSSPPITPPASETKMSPRKALIPVSKSQSAEACSESRNRVRRDSSCL 560
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 365 LRNWTRTPSSPPITPPASETKISSPRKALIPVSKSQSAEACSESRNRVRRDSSCL 424
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 561 ESKVKQKVCNCVTELDGQVENLHLDLCLAGNOEDLSKDSLGPTKSKIEGAGTSISE 620
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 425 ESKVKQKVCNCVTELDGQVAESLRLDCLCLSGTQEVLSQDSGPTKSKIEGAGTSISE 484
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 621 PPSPISPVASESCGTLPLPLPCCGEGSEWGVKENSPPNKWLLAMAARKAENSPRSP 680
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 485 PPSPISPVASESCGTLPLPLPCCGEGSEWGVKENSPPNKWLLAIAARKAENSPRSP 544
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 681 SSOTNSRRSQSGKTLPSVTTIPSSMRKICTYFHRKQEDFCGPEHSTEL 730
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 545 SSOTNSRRSQSGKTLPSVTTIPSSMRKICTYFHRKQEDFCGPEHSTEL 594
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
Q9NW03_HUMAN PRELIMINARY; PRT; 434 AA.
AC Q9NW03;
CT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-MAR-2006, entry version 24.
DE CDNA FLJ10399 fis, clone NT2RM400354, weakly similar to
DE LETHAL(2) DENTICLELESS PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14702039; DOI=10.1038/ngl1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagabari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
```

RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Inose N.,
RA Musashino K., Yuuki H., Ohshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii K., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi O., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45 (2004).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AK001261; BAA91586.1; -; mRNA.
DR Ensembl; ENSG00000143476; Homo sapiens.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 2.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 2.
DR PROSITE; PS00082; WD_REPEATS_2; 2.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 434 AA; 46677 MW; 1B73E2BF1311155D CRC64;

Query Match 59.1%; Score 2293; DB 2; Length 434;
Best Local Similarity 100.0%; Pred. No. 7.7e-119;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 297 MFNMTGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDRAAYIKWVSTPWPPTVL 356
DB 1 MFNMTGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDRAAYIKWVSTPWPPTVL 60

QY 357 LGHSQEVTSVCWCPSPDTKIATCSDNTLKIWLNRGLEKPGDKLSTVGWASQKKKES 416
DB 61 LGHSQEVTSVCWCPSPDTKIATCSDNTLKIWLNRGLEKPGDKLSTVGWASQKKKES 120

QY 417 RPLGLVTVTSQSSTPAKAPRVKCNPSNPSSSAACPSCAGDLPLPSNTPFTSIKTSPPAKA 476
DB 121 RPLGLVTVTSQSSTPAKAPRVKCNPSNPSSSAACPSCAGDLPLPSNTPFTSIKTSPPAKA 180

QY 477 RSPINRSGSVSSVPKPPSPFKMSIRNWTRTPSSPPITPPASETKIMSPKALIPVSQ 536
DB 181 RSPINRSGSVSSVPKPPSPFKMSIRNWTRTPSSPPITPPASETKIMSPKALIPVSQ 240

QY 537 KSSQAEACSSSRNRVRRLDSSCLESVKQKVCNVCVTLDGOVENLHLDLCLAGNOE 596
DB 241 KSSQAEACSSSRNRVRRLDSSCLESVKQKVCNVCVTLDGOVENLHLDLCLAGNOE 300

QY 597 DLSKDSLGPTKSKIEGAGTSISEPPSPISPYASESGCTLPLPLPCGEGSEMVGKENS 656
DB 301 DLSKDSLGPTKSKIEGAGTSISEPPSPISPYASESGCTLPLPLPCGEGSEMVGKENS 360

QY 657 PENKNWLLAAAKRKAENPSRPSSTQTPNSRQSGKTLPSPTVITPSSMRKICTVFHRK 716
DB 361 PENKNWLLAAAKRKAENPSRPSSTQTPNSRQSGKTLPSPTVITPSSMRKICTVFHRK 420

QY 717 SQEDFCGPEHSTEL 730
DB 421 SQEDFCGPEHSTEL 434

RESULT 11
05ZJW8_CHICK PRELIMINARY; PRT; 720 AA.
AC Q5ZJW8;
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE Hypothetical protein.
DE ORFNames=RCJMB04.15a2;
GN Gallus gallus (Chicken).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT gene function analysis.";
RL Genome Biol. 6:R6-R6(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AJ270316; CAG31975.1; -; mRNA.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 2.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS00082; WD_REPEATS_2; 4.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 720 AA; 78633 MW; 1B9101FCA671EB4E CRC64;

Query Match 56.9%; Score 2206.5; DB 2; Length 720;
Best Local Similarity 61.9%; Pred. No. 8.8e-114;
Matches 447; Conservative 73; Mismatches 161; Indels 41; Gaps 12;

QY 20 SSQYPLQSLITCVCQSGNDEHTSYGETGVPPGCTESSAPNMEHVLAVANEEGVRLY 79
DB 16 SSSLPLQLHLDGYRCSREDDHLSYGEIMGVPPFGCSAANFEHVLAVANEEGVRLY 75

QY 80 NTESOSFRKKCFKEMAHMNAVFDLAWVPEGLKLVTAAGDQAKFWDVKAGELIGTCCKG 139
DB 76 DTEAQTTLKISKQWQAHNSNAVFDLAWVPEHRIVTASGDQAKVWDVZAGELLGICKGH 135

QY 140 QCSLKSVAFSKFEKAVFTCTGDRGNIMVMDTRCNKKDGYRQVQNIISGAHNTSDKQTPSK 199
DB 136 QCSLKSVAFSKFEKAVFTCTGDRGNIMVMDTRCNKKDGYRQVQNIISGAHNVDRQTPSK 195

QY 200 -PKKQNSKGLAPSVDFQOSVTVVLFDENTLVAGAVDGIKVDLKNRYTAYRQEP 258
DB 196 LRKCKRQLRGLAPLVDFOQSVTVLLQDHTLISAGVDGVIKVDLKNRYAAYRQDPVP 255

QY 259 SKSFLYPGSSSTKGLGYSSLLDSTGLTGFANCTDDNIYMFNMTGLKTSVPAIFNGHONST 318
DB 256 SKSFFYPGTSRKLGYSSLLDSTGLTGFANCTDDSIYMFNMTGLTFFVAVFSGHONST 315

QY 319 FYVKSLSLSPDDQFLVSGSSDEAAYIKWVSTPWPPTVLVLGHSQEVTSVCWCPSDFTKIAT 378
DB 316 FYIKSSISLDDQFLVSGSSDCNAYIKWVSEPSLPPRLVGHSEVTSIACPSDFTKIAT 375

QY 379 CSDNTLKIWLNR -GLEKPGDKLSTVGWASQKKESRPGLVTVTSQSSTPAKAPRVK 437
DB 376 CSDNTVIRIWLQHYPEEKSVSNKAKLVGWVTKKPEEQRQAGRSASQSTPAKAFSVG 435

QY 438 CNPSNSSPSSAACAPSCAGDLPLPSNTPFTSIKTSPPAKARSPINRSGSVSSVPKPPSS 497

Db 77 VTIFNTGEKQS---SVLKEWQAHDNAVFDAIWAUVPGTNCLVWTASGQOTARLWDVITGDLLG 133
Qy 135 TCKGHQCSLSKVAFSKFEKAVFCTGGRDGNIMVWDTKCNKKGDFYQVNOISGAHNTSDK 194
Db 134 TFKGHQCSLSKVAFYKOEKAVFSTGGRDGNIMIWDTRCSKKGDFYQVNOISGAHNTSDK 193
Qy 195 QTPSKPKKQNSKGLAPSVDFOQSVTVVLFDQDENTLVASAGAVDGIKWDLRKNYAYRQ 254
Db 194 FFPQTKR---RGMAPPVDSQGVTVLFCDETKLISSGAVDGIKWDLRKNYAYRQ 249
Qy 255 EPIASKSFLYPGSSTRKGLYSLIISTGSTFLPANCDDNIMYFNMGTGLKSPVAIFNGH 314
Db 250 NPLPLQAYPYPGSCTRKGLYSLIISTGSTFLPANCDDNIMYFNMGTGLKSPVAIFNGH 309
Qy 315 QNSTFVVKSLSDDDQFLVSGSSDEAAYIWKVSTPMQPTVLIGHSEVTSVCWCPSDFT 374
Db 310 SNSSFYKSVSPDDQFLASGSSDHNVIWKISDPKQAPMMLQGHSEVTSVAVWCPTDFT 369
Qy 375 KIATCSDDNTLKIWRNLNRGLEEKPGGDKL-----STVGWASQKKESR--PGLVTVTSQ 427
Db 370 KIATCSDDNTVRIWNLNR---KPEGENSTIQDGNLVGWTIRKQVSPNRTPG--HHSPE 423
Qy 428 STPAKAPRVKCNPSNSPSSAACAPSCAGDLPLPSNTPFTSITSPAKARSPINRGVS 487
Db 424 LTPSKNPGSVRSVSLASPOPATCAPTCAA-LPLPSNTSS---APPAKL-----467
Qy 488 SVSPKPPSPFKMSIRNWTPTSSSPITPPASETKIMSPKALIPVSKSSQAEACSES 547
Db 468 -TSPKMPSS---SLOQWISR--SSKSPV-----RKALTPVLOGLS-----F 501
Qy 548 RNRVRRRLDSSCLESVKQKVCNCVTELDQOVENLHLDLCLAGNQEDLSKDSLGPTK 607
Db 502 EHRVXRLETG-----DSASSGLG---520
Qy 608 SSKIEGAG---TSISEPPGPISPYASESCOTLPLRPGEGSEVMYKENSOPENKNWLL 664
Db 521 -EIDGVSELYPNVKRSRVSSTLKKEDSFGLESEKRLGSDGAEASGKENSPPRTDWLS 579
Qy 665 AMNAKKAENPSRPSSTQTPNSRRQSGKTLPSVITP-----SSMR 707
Db 580 VISQKFKG-SAQPKSPSS---GSSQODTRTLESAAVSPRPMKVFPPTNKKASPSKPMK 635
Qy 708 KICTYFHRKSOE 719
Db 636 KISSYFWKRTQD 647

Search completed: November 8, 2006, 18:13:51
Job time : 315 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:15:55 ; Search time 44 Seconds
(without alignments)
1408.400 Million cell updates/sec

Title: US-10-726-160-2

Perfect score: 3877

Sequence: 1 MLFNSVLROPQLGVLRNGWS.....TYFHRKSQEDFCGPEHSTEL 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 326800 seqs, 84889852 residues

Total number of hits satisfying chosen parameters: 326800

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New.*

- 1: /EMC_Celerra_SID33/ptodata/1/pubpaa/US09_NEW_PUB.pap.*
- 2: /EMC_Celerra_SID33/ptodata/1/pubpaa/US06_NEW_PUB.pap.*
- 3: /EMC_Celerra_SID33/ptodata/1/pubpaa/US07_NEW_PUB.pap.*
- 4: /EMC_Celerra_SID33/ptodata/1/pubpaa/US08_NEW_PUB.pap.*
- 5: /EMC_Celerra_SID33/ptodata/1/pubpaa/PCT_NEW_PUB.pap.*
- 6: /EMC_Celerra_SID33/ptodata/1/pubpaa/US10_NEW_PUB.pap.*
- 7: /EMC_Celerra_SID33/ptodata/1/pubpaa/US11_NEW_PUB.pap.*
- 8: /EMC_Celerra_SID33/ptodata/1/pubpaa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3503.5	90.4	673	6	US-10-540-898-674
2	2914.5	75.2	784	6	US-10-540-898-671
3	759	19.6	146	6	US-10-540-898-676
4	267	6.9	321	6	US-10-953-349-33327
5	255	6.6	278	6	US-10-953-349-33329
6	255	6.6	279	6	US-10-953-349-33328
7	250	6.4	415	7	US-11-293-697-4324
8	249	6.4	301	6	US-10-449-902-36351
9	241	6.2	343	6	US-10-449-902-56513
10	240	6.2	244	6	US-10-449-902-41812
11	234	6.0	317	7	US-11-056-355B-90488
12	234	6.0	317	7	US-11-056-355B-94244
13	231.5	6.0	319	7	US-11-056-355B-13097
14	231.5	6.0	346	7	US-11-056-355B-13096
15	231.5	6.0	1281	6	US-10-449-902-42332
16	227	5.9	808	6	US-10-449-902-41350
17	219	5.6	324	6	US-10-449-902-36608
18	217.5	5.6	299	7	US-11-056-355B-23167
19	217.5	5.6	299	7	US-11-056-355B-77008
20	217.5	5.6	299	7	US-11-056-355B-79227
21	217.5	5.6	302	7	US-11-056-355B-84740
22	217.5	5.6	302	7	US-11-056-355B-77007
23	217.5	5.6	304	7	US-11-056-355B-23166
24	217.5	5.6	304	7	US-11-056-355B-84739
25	217.5	5.6	837	7	US-11-056-355B-108275

26	217.5	5.6	837	7	US-11-056-355B-119514	Sequence 119514, A
27	217	5.6	573	6	US-10-449-902-51734	Sequence 51734, A
28	210.5	5.4	803	7	US-11-056-355B-88301	Sequence 88301, A
29	210.5	5.4	803	7	US-11-056-355B-92057	Sequence 92057, A
30	205.5	5.3	247	6	US-10-449-902-54161	Sequence 54161, A
31	204	5.3	680	6	US-10-449-902-45431	Sequence 45431, A
32	202.5	5.2	299	6	US-10-953-349-16820	Sequence 16820, A
33	202.5	5.2	299	7	US-11-056-355B-59380	Sequence 59380, A
34	202.5	5.2	314	6	US-10-953-349-16819	Sequence 16819, A
35	202.5	5.2	314	7	US-11-056-355B-59379	Sequence 59379, A
36	201.5	5.2	267	7	US-11-056-355B-37531	Sequence 37531, A
37	201.5	5.2	267	7	US-11-056-355B-85524	Sequence 85524, A
38	201.5	5.2	333	7	US-11-056-355B-37530	Sequence 37530, A
39	201.5	5.2	333	7	US-11-056-355B-85523	Sequence 85523, A
40	201.5	5.2	891	6	US-10-449-902-54504	Sequence 54504, A
41	201.5	5.2	922	7	US-11-056-355B-75239	Sequence 75239, A
42	201	5.2	353	6	US-10-449-902-40839	Sequence 40839, A
43	199	5.1	2629	6	US-10-547-530-12	Sequence 12, Appl
44	196	5.1	874	7	US-11-056-355B-78161	Sequence 78161, A
45	192	5.0	246	6	US-10-449-902-50240	Sequence 50240, A

ALIGNMENTS

RESULT 1

US-10-540-898-674
; Sequence 674, Application US/10540898
; Publication No. US20060166213A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)
; CURRENT APPLICATION NUMBER: US/10/540,898
; PRIOR FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: US 10/330,773
; PRIOR FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 674
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-540-898-674

Query Match 90.4%; Score 3503.5; DB 6; Length 673;
Best Local Similarity 91.8%; Pred. No. 1e-185;
Matches 670; Conservative 1; Mismatches 2; Indels 57; Gaps 2;

QY	1	MLFNSVLROPQLGVLRNGWSQYPIQSLITGYQCSGNDHSTSYGFTGVPPPGCTFSSA	60
DB	1	MLFNSVLROPQLGVLRNGWSQYPIQSLITGYQCSGNDHSTSYGFTGVPPPGCTFSSA	60
QY	61	PNMEHVLAVANEEGFVRLYNTESQSFRRKCKPEWMAHNAVFDLAWVPGLKLVTAAGQ	120
DB	61	PNMEHVLAVANEEGFVRLYNTESQSFRRKCKPEWMAHNAVFDLAWVPGLKLVTAAGQ	120
QY	121	TAKFMDVKAGELIGTCKGHQCSLKSVAFSKPKAVFCTGGRDGNIMVWOTRCKNKDGFYR	180
DB	121	TAKFMDVKAGELIGTCKGHQCSLKSVAFSKPKAVFCTGGRDGNIMVWOTRCKNKDGFYR	180
QY	181	QVNIQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQOSVTVVLFQDENTLVSGAVDGI	240
DB	157	-----DFQQSVTVVLFQDENTLVSGAVDGI	183
QY	241	KVWDLRKNVTAYRORPIASKSFLYPGSSRKLKGLYSSLLDSTGLFANCTDDNIYFNM	300
DB	184	KVWDLRKNVTAYRORPIASKSFLYPGSSRKLKGLYSSLLDSTGLFANCTDDNIYFNM	243
QY	301	TGLKTSFPAIFNGHQNSTFYVKSLSLSPDQFVLVSGSSDEAAIYKIVSTPWQPTVLLGHS	360
DB	244	TGLKTSFPAIFNGHQNSTFYVKSLSLSPDQFVLVSGSSDEAAIYKIVSTPWQPTVLLGHS	303

Db 61 CLESVQKCVKNCVTELDGQVENHLHLDCCLAGNQEDLSKSLGPTKSKKEGAGTSI 120
QY 619 SEPPSPISPYASRSCGTLPLRPC 643
Db 121 SEPPSPISPYASRSCGTLPLRPC 145

RESULT 4
US-10-953-349-33327
; Sequence 33327, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33327
; LENGTH: 321
; TYPE: PRP
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33327

Query Match 6.9%; Score 267; DB 6; Length 321;
Best Local Similarity 24.7%; Pred. No. 1.2e-07;
Matches 98; Conservative 61; Mismatches 140; Indels 98; Gaps 15;
QY 58 SSAPNMEHLAVANESEGFVLYNTEQSFRRKCKFKEMAHWNAVFDLAWVPGLKLVATA 117
Db 11 SASNDKSAIIWEDEDEGLLRHLTL-----GHRKSVMMVWSPDDCQLLTG 58
QY 118 GDQAKFWDYKAGELICTCKGHCQSLKSVAFSKFAVFCCTGGRDGNIMVWTRCNKXDG 177
Db 59 QEBETIRRDVKSGLHVKSVKATGLISCW--FPDGKQLSGLADQNFQCIWDLGKEVDC 116
QY 178 FYRQVNIQSAHNTSKDTPSK-----PKKQNSKGLAPSVDFQVQSVTVLFDENTLVS 232
Db 117 WKQ-----GSTGSDFAVAKDGNLIISMKN-----TILLFDRET----- 153
QY 233 AGAVDGLIKVLDLRKNVAVRQSPISKSLYFGSSTRKGLYSSLIIDSTGSLFANCTD 292
Db 154 -----KQERIIE-----GST-----ITSFCLSEDGDFLLVNLVS 183
QY 293 DNIYFMNTGLKTSPLVAI--FNHQNSTFYVKSLSLSPDDQ-FLVSGSDEAAIYIKVSTP 349
Db 184 EEIHLWN---IRNDPVRVNRVNGHKTRFVIRSCFGFQAIFASGSDSQVYIHRATG 240
QY 350 WQPTVLLGHSEVTSVCWCPDFTKIATCSDNTLKIWLNRGLBEKPGDKLSTVGWA 409
Db 241 DLIET-LAGHSGTVNCVSWNPANPHMLASASDDHTVRIW-----GAKK-----S 283
QY 410 SOKKESRPLVTVTSQSPTAKAPRV---KCNPSNS 443
Db 284 SLKRKDVGCSSNCNSRSGNNAHNVFVQKCGNNT 320

RESULT 5
US-10-953-349-33329
; Sequence 33329, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 33329
; LENGTH: 278
; TYPE: PRP
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33329
Query Match 6.6%; Score 255; DB 6; Length 278;
Best Local Similarity 25.6%; Pred. No. 4.5e-07;
Matches 90; Conservative 51; Mismatches 124; Indels 86; Gaps 14;
QY 104 LAWVPGLKLVTAAGDQTAKFWDVYKAGELICTCKGHCQSLKSVAFSKFAVFCCTGGRDG 163
Db 2 VAWSPDDCQLLTGQEBETIRRDVKSGLHVKSVKATGLISCW--FPDGKQLSGLADQ 59
QY 164 NIMVWTRCNKXDGFYRQVNIQSAHNTSKDTPSK-----PKKQNSKGLAPSVDFQOS 218
Db 60 NFCIWDLDGKEVDCWKQ-----GSTGSDFAVAKDGNLIISMKN----- 101
QY 219 VTVLFDQENTLVSAGAVDGLIKVLDLRKNVAYRQEPISKSLYFGSSTRKGLYSSLI 278
Db 102 -TILLFDRET-----KQERIIE-----GST-----ITSFC 136
QY 279 LDSTGSLFANCTDNIYFMNTGLKTSPLVAI--FNHQNSTFYVKSLSLSPDDQ-FLVSG 335
Db 127 LSEDGDFLLVNLVSEIHLWN---IRNDPVRVNRVNGHKTRFVIRSCFGFQAIFASG 183
QY 336 SDEAAIYIKVSTPWPQPTVLLGHSEVTSVCWCPDFTKIATCSDNTLKIWLNRGLE 395
Db 184 SEDSQYIWHRATGDIET-LAGHSGTVNCVSWNPANPHMLASASDDHTVRIW----- 235
QY 396 EKPGGDKLSTVGWASOKKESRPLVTVTSQSPTAKAPRV---KCNPSNS 443
Db 236 ---GAKK-----SLKRKDVGCSSNCNSRSGNNAHNVFVQKCGNNT 277

RESULT 6
US-10-953-349-33328
; Sequence 33328, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33328
; LENGTH: 279
; TYPE: PRP
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33328

Query Match 6.6%; Score 255; DB 6; Length 279;
Best Local Similarity 25.6%; Pred. No. 4.5e-07;
Matches 90; Conservative 51; Mismatches 124; Indels 86; Gaps 14;
QY 104 LAWVPGLKLVTAAGDQTAKFWDVYKAGELICTCKGHCQSLKSVAFSKFAVFCCTGGRDG 163
Db 3 VAWSPDDCQLLTGQEBETIRRDVKSGLHVKSVKATGLISCW--FPDGKQLSGLADQ 60
QY 164 NIMVWTRCNKXDGFYRQVNIQSAHNTSKDTPSK-----PKKQNSKGLAPSVDFQOS 218
Db 61 NFCIWDLDGKEVDCWKQ-----GSTGSDFAVAKDGNLIISMKN----- 102
QY 219 VTVLFDQENTLVSAGAVDGLIKVLDLRKNVAYRQEPISKSLYFGSSTRKGLYSSLI 278
Db 103 -TILLFDRET-----KQERIIE-----GST-----ITSFC 127
QY 279 LDSTGSLFANCTDNIYFMNTGLKTSPLVAI--FNHQNSTFYVKSLSLSPDDQ-FLVSG 335
Db 128 LSEDGDFLLVNLVSEIHLWN---IRNDPVRVNRVNGHKTRFVIRSCFGFQAIFASG 184

QY 336 SDEAAYIKWSTPWPPTVLGHSEVTSVCWPSDFTKIATCSDNTLKIWLNRGL 395
DB 185 SEDSQYIWHRTAGDIET-LAGHSGTNCVSWNFPANPHMLASADDDHTVRIW- 236
QY 396 EKPGGDKLSTVGWASQKKESRPLVTVTSSOSTPAKAPRV---KCNPSNS 443
DB 237 ---GAKK-----SSLKRDVSSNCNSNGSRSGNAHNHNVFQKNGNNT 278

RESULT 7

US-11-293-697-4324
; Sequence 4324, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4324
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4324

Query Match 6.4%; Score 250; DB 7; Length 415;
Best Local Similarity 24.3%; Pred. No. 1.3e-06;
Matches 89; Conservative 58; Mismatches 135; Indels 84; Gaps 14;

QY 35 SGNDEHTSGTGVVPV-----PFGCTFSSAPNMEHVLAVANEGBGVRLYNTESQSRKK 89
DB 123 SBEELNTLEGRNVVYAIAPNPYG-----DKIATGSDTKLWSVET-----GK 168
QY 90 CFKEMAHNAVDFDLAWPGEIKLVTAAGDQTAKEFMDVAVKAGELIGTCCKHQSLSKSVAPS 149
DB 169 CYHTPRGTAETVCLSFNQSLTAVTGMSTAKLWDIQNGEVEVTLRGHSAEIIISLSE 228
QY 150 KPEKAVCTCGRGNIMVMDTRCNKDGQYRQVQNGISGAHTSDKQTPSPKPKKQNSKGL 209
DB 229 TSGDRII-TGSDFTVTVVVDATGKR-----VNLIIG----- 259
QY 210 APSVDQQSVTVVLFQDENTLVSAGAVDGIIVKWLKKNYATYRQPIASKSFLYPGSST 269
DB 260 -----HCAEISSAFWDCSLILTGSMDTKLWD-----ATNGKCVATLT----- 300
QY 270 RKLGYSSLILDS-----TGSTLFANCTDDNIYFMNMTGLKTSVPAIFNGHQNSTFFVYKSSL 325
DB 301 ---GHDDEILDSCFDVTGKLIATASADGTARIFSAATKRC--IAKLEGEHEG--ISKISF 353
QY 326 SPDQQLVSGSDAAYIKWSTPWPPTVLGHSEVTSVCWPSDF--TKIATCSDDN 383
DB 354 NFGQNHLLTSGSDKTARINDAQT-GQCLQVLEGTDEIFS---CAPNYKGNIVITGSKDN 409
QY 384 TLKIWR 389
DB 410 TCRIMR 415

RESULT 8

US-10-449-902-36351
; Sequence 36351, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36351
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-36351

Query Match 6.4%; Score 249; DB 6; Length 301;
Best Local Similarity 25.4%; Pred. No. 1.1e-06;
Matches 86; Conservative 51; Mismatches 124; Indels 78; Gaps 13;
QY 66 VLAVA-NEEG-----FVRLYNTESQSRKKCFKEMAHNAVDFDLAWPGEIKLV 114
DB 24 VLAVFRNDRGNYCLSCGKDIRIIRLWNPHTGAL----VKPYKSHGREVDVNSSSDNAKLV 79
QY 115 TAAGDQTAKEFMDVAVKAGELIGTCCKHQSLSKSVAFSKFAVCTCGRGNIMVMDTRCNK 174
DB 80 SCGDRQVYWDVASARVIRKFRGHSEINSEVKNF--NTVVNSAGYDRTVRAFDCR--- 135
QY 175 KDFYRQVQNGISGAHTSDKQTPSPKPKKQNSKGLAPSVDFQSVTVVLFQDENTLVSAG 234
DB 136 -----SQSSDPIQTIDTFQDSVMSVNL--TNTETIAG 165
QY 235 AVDGLIKWDLKKNYATYRQPIASKSFLYPGSSTRLKGY--SSLILDSTGTLFANCTD 292
DB 166 SVDTGRTIFDIMG-----RETVDNLGHPVNCISLSDNRNCLLANCLD 208
QY 293 DNIYMFN-MTGLKTSVPAIFNGHQNSTFYKSSLSPPDQFLVSGSSDEAAIWKVSTPWQ 351
DB 209 STVRLDKSTG---ELLOEYKGHICKSPKMDCCLTNDDAFVVGSGSEDGYIFFWELV---D 262
QY 352 PPTV--LLGHSEVTSVCWPSDFTKIATCSDNTLKIW 388
DB 263 APVSSPRAHSSVTVTSVYHPTACML-TSSVDGTIRVM 300

RESULT 9

US-10-449-902-56513
; Sequence 56513, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56513
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-56513

Query Match 6.2%; Score 241; DB 6; Length 343;
Best Local Similarity 23.4%; Pred. No. 3.4e-06;
Matches 74; Conservative 52; Mismatches 114; Indels 76; Gaps 10;

QY 92 KEMW---AHNNAVFDLAWPGLKLVTAAGDQTAKEFDWVKAGBLIGTCCKGHCQSLKSVAF 148
Db 87 KNYMVLRGHKNVLDLQWITDGTQIISASPDKTVRVWDVETGKQVKMAEHSSFFVNSCCP 146
QY 149 SKEKAVFTCGGRGNIMWDTTCNKKDGFYRQVQNSGAHNTSDKQTPSKPKKQNSKG 208
Db 147 ARKWPLVVGSGDGTAKLWDLR-----QRG 172
QY 209 LAPSVDFQOSVTVLVFQDENTLVLSAGAVDGLIKVWDLRKN-VTAYRQEPISKSLFPLPGS 267
Db 173 AIGLTPDKVQITAVSSEAADKVTGGLDNDVWDLRKNNEVTEYDK-----219
QY 268 STRKLGYSLLI-----LDSTGSTLFANCTDDNIWMFN-----TGLKTSPPVAFNGHQN 316
Db 220 -----GHQDMITCQWLSPOGSSYLLTNAMONELKIWDLRPVAPENRIKT-----LTGHQH 269
QY 317 S--TFVVKSLSPDDQLVSGSDDEAAYIKWSTPQPPVTLLGHGQEVTSVCWCPSDFT 374
Db 270 NFEKNLLKCSWSPDNKRVTAGSADRMVYIWD-TTSRRILYKLPFHNGSVNETAFHPTE-P 327
QY 375 KIATCSDDNTLKTWRL 390
Db 328 VIGSCGSDKQIYLDEL 343

RESULT 10
US-10-449-902-41812
; Sequence 41812, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 41812
; LENGTH: 244
; TYPE: prt
; ORGANISM: Oryza sativa
US-10-449-902-41812

Query Match 6.2%; Score 240; DB 6; Length 244;
Best Local Similarity 29.1%; Pred. No. 2.6e-06;
Matches 65; Conservative 36; Mismatches 86; Indels 36; Gaps 6;

QY 307 PVAIFNGHONSTFYVKSSLPDQFLVSGSDDEAAYIKWSTPQPPVTLLGHGQEVTSV 366
Db 5 PIKAYTGSKTESFVKSAISPDGTHILGSGSDGNVILWQVDPQERGPILLEGHEGATS 64
QY 367 CWCPSDFTKIATCDDNTLKIWLNRGLEKPGGDKLSTVGWASQKKESRPLGLVTTSS 426
Db 65 DWCASEVGKIATSDSDSKVRVMTERRV--FPNTSSPTVI-----RKRITAPN----TGS 113
QY 427 QSTPAKAPRVKCNPSNSPSSAACAGDLPLPSNP-----TFSIKTSPAK-----475
Db 114 RSASHEL-----ATTSRDYGVAACTSDAGELPTGSRSPQLRVLEFGTPEAKGAFRL 167
QY 476 -----ARSPNRGVSVSVPKPPSKMIRNVTTPS 510
Db 168 FQEDSLDIRKSPQAQMNPSVSPPHSLKRRITRDYFASSSS 210

RESULT 11
US-11-056-355B-90488

; Sequence 90488, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 90488
; LENGTH: 317
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; LOCATION: (1)..(317)
; OTHER INFORMATION: Ceres Seq. ID no. 12724103
US-11-056-355B-90488

Query Match 6.0%; Score 234; DB 7; Length 317;
Best Local Similarity 22.7%; Pred. No. 7.5e-06;
Matches 75; Conservative 66; Mismatches 126; Indels 64; Gaps 11;

QY 66 VLAVANEGBFVLY--NTESQSFRRKCKPEKMAHNAVFDLAWPGLKLVTAAGDOTAK 123
Db 38 LLASASADKTIRTYTINTINDPIAP-VQEFYTGHEGHSIDVAFSSDARFIVASDDKTUK 96
QY 124 FWDVKAAGELIGTCCKGHCQSLKSVAFSKPEKAVFTCGGRGNIMWDMTRCNKKDGFYRQV 183
Db 97 LMDVETGSLIKTLIGHTNYAFVCNFP-QSNMIVSGSPDETVRINDVTTGK-----CU 148
QY 184 QISGAHNTSDKQTPSKPKKQNSKGLAPSDVPQQSVTVLVFQDENTLVLSAGAVDGLIKW 243
Db 149 KVLPAHS-----DPTAVDFNRDGLSVSSSYDGLCRIW 182
QY 244 D-----LRKNYTAQRQEPISKSLFPLPGSSTRKLGYSLLDSTGTLFANCTDDNIYM 298
Db 183 DSGTGHCVKTLIDDPNPPVSVRF-----SPNGKFILVGLTLDNTLRW 225
QY 299 NMTGLKTSPPVAFNGHONSTFYVKSSLS-PDQFLVSGSDDEAAYIKWSTPQPPVTLL 357
Db 226 NISSAKF--LKTYTGHVNAQYCISSAFSVTNGKRVISGSDNCVHWELNSK-KLLQKLE 282
QY 358 GHSQEVTSVCWCPDFTKIATCDDNTLKIW 388
Db 283 GHTETVMNVACHPTE-NLIASGSLDKTVRIW 312

RESULT 12
US-11-056-355B-94244
; Sequence 94244, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 94244
; LENGTH: 317
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide

```
; LOCATION: (1)..(317)
; OTHER INFORMATION: Ceres Seq. ID no. 12724103
US-11-056-355B-94244

Query Match
Best Local Similarity 6.0%; Score 234; DB 7; Length 317;
Matches 75; Conservative 66; Mismatches 126; Indels 64; Gaps 11;

QY 66 VLAVANEQFVRLY--NTSQSFRKCKFKEWMAHNAVFDLAWVPGELKLVTAAGDQTA 123
DB 38 LLASASADKTIIRYTTINTNDPIAEP-VQEFTHENGISDVAFSSDARFVBSADKTLK 96
QY 124 FWDVAGELIGCKHQCSLKSVAFSKFAVCTGGRDGNIMWDTTCNKKDGFYRQVN 183
DB 97 LMDVETGSLIKTLIGHTNYAFVNFNP-QSNMIVSGSFDETIRIMDVTTGK-----CL 148
QY 184 QISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSVTVVLFQDENTLVSAGAVDGI 243
DB 149 KVLPAHS-----DPTAVDFNRDGLSVSSVSDGLCRIM 182
QY 244 D-----LRKNYATAYROEPIASKSFLYPGSSTRKLGYSLLDSTGLTFANCTD 298
DB 183 DSGTGHCVKTLIDENPPVSVRF-----SPNGKFIIVGTLDNTLRW 225
QY 299 NMTGLKTSVPAIFNGHONSTFYVKSSLS-PDQFLVSGSSDEAAIYIKVSTPWQPT 357
DB 226 NISSAKF--LKTYTGHVNAQYCISSAFSVTNGKRIVSGSEDNCVHMWELNSK-K 282
QY 358 GHSQEVTSVCMCPDFTKIATCDDNTLKIW 388
DB 283 GHTETVMNVACHPTE-NLIASGLDKTVRIW 312

RESULT 13
US-11-056-355B-13097
; Sequence 13097, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: 2005-02-14
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 13097
; LENGTH: 319
; TYPE: prt
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(319)
; OTHER INFORMATION: Ceres Seq. ID no. 12344627
US-11-056-355B-13097

Query Match
Best Local Similarity 6.0%; Score 231.5; DB 7; Length 319;
Matches 84; Conservative 63; Mismatches 142; Indels 93; Gaps 14;

QY 20 SSOYPLQSLITGYQCSGNDHTSYGTGVPVPPFGCTFSAPNMEHVLAVANEQFVRLY 79
DB 12 SPGYVLRSTLEGR-----RAVSTVKFSPDG-----RLLASASADKLLRW 52
QY 80 NTESQSRKCKFKEWMAHNAVFDLAWVPGELKLVTAAGDQTA 136
DB 53 SSSDLT-----PVAELGHEGEGVSDLSFSPDGRLLASASDDRTVRIWDLAVGGARLVK 108
QY 137 KGHQCSLKSVAFSKFAVCTGGRDGNIMWDTTCNKKDGFYRQVNQISGAHNTSDK 196
DB 109 TGHNTYAFVCSFSPHGN-VLASGSFDETIRVWEVRSKG-----CLRVLPAHS----- 154
```

```
QY 197 PSKPKKQNSKGLAPSVDFQSVTVVLFQDENTLVSAGAVDGI 251
DB 155 -----EPVTAVDFDREGDMIVSGSYDGLCKRWSDSAGCHCVKTLID 194
QY 252 YRQEPPIASKSFLYPGSSTRKLGYSLLDSTGLTFANCTDDNIYMFNMTG---LKTSPV 308
DB 195 DESPPVSPAKF-----SPNGKFIILATLDTLRLWNFSAKGFLKT--- 234
QY 309 AIFNGHONSTFYVKSSLS-PDQFLVSGSSDEAAIYIKVSTPWQPTVLLGHSQEVTSV- 366
DB 235 --YTHVNTKVCIPAAFSITNSKYIVSGSEDKCVYLDLQSR-RIVQKLEGHTDITVIAVS 291
QY 367 CWCPSDFTKIATCDDNTLKIW 388
DB 292 CHPKENMIASGALDNDKTVKM 313

RESULT 14
US-11-056-355B-13096
; Sequence 13096, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: 2005-02-14
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 13096
; LENGTH: 346
; TYPE: prt
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(346)
; OTHER INFORMATION: Ceres Seq. ID no. 12344626
US-11-056-355B-13096

Query Match
Best Local Similarity 6.0%; Score 231.5; DB 7; Length 346;
Matches 84; Conservative 63; Mismatches 142; Indels 93; Gaps 14;

QY 20 SSOYPLQSLITGYQCSGNDHTSYGTGVPVPPFGCTFSAPNMEHVLAVANEQFVRLY 79
DB 39 SPGYVLRSTLEGR-----RAVSTVKFSPDG-----RLLASASADKLLRW 79
QY 80 NTESQSRKCKFKEWMAHNAVFDLAWVPGELKLVTAAGDQTA 136
DB 80 SSSDLT-----PVAELGHEGEGVSDLSFSPDGRLLASASDDRTVRIWDLAVGGARLVK 135
QY 137 KGHQCSLKSVAFSKFAVCTGGRDGNIMWDTTCNKKDGFYRQVNQISGAHNTSDK 196
DB 136 TGHNTYAFVCSFSPHGN-VLASGSFDETIRVWEVRSKG-----CLRVLPAHS----- 181
QY 197 PSKPKKQNSKGLAPSVDFQSVTVVLFQDENTLVSAGAVDGI 251
DB 182 -----EPVTAVDFDREGDMIVSGSYDGLCKRWSDSAGCHCVKTLID 221
QY 252 YRQEPPIASKSFLYPGSSTRKLGYSLLDSTGLTFANCTDDNIYMFNMTG---LKTSPV 308
DB 222 DESPPVSPAKF-----SPNGKFIILATLDTLRLWNFSAKGFLKT--- 261
QY 309 AIFNGHONSTFYVKSSLS-PDQFLVSGSSDEAAIYIKVSTPWQPTVLLGHSQEVTSV- 366
DB 262 --YTHVNTKVCIPAAFSITNSKYIVSGSEDKCVYLDLQSR-RIVQKLEGHTDITVIAVS 318
QY 367 CWCPSDFTKIATCDDNTLKIW 388
```

Db 319 CHPKENNIASGALDNDRKTKVW 340

RESULT 15
US-10-449-902-42332
; Sequence 42332, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42332
; LENGTH: 1281
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-42332

Query Match 6.0%; Score 231.5; DB 6; Length 1281;
Best Local Similarity 23.8%; Pred. No. 5.2e-05;
Matches 98; Conservative 60; Mismatches 143; Indels 111; Gaps 20;

QY 72 EEGFVRLYNTESQSRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQTAKEFMDVKAGR 131
Db 625 QEGRVWVEWNRVEYLSPTQTLRGHTGPGVFGTYSADGQRLVSGSSDLSVRVWDASTGE 684

QY 132 LIGTCKGHCQSLKSAFSEKFAVFTCTGGRDGNIMVMDT---RCNKK---DGFYRQV-- 182
Db 685 SLRELKGTADIRAVALSADGQRI-ASSGDDQTVRVWDASTGECLELKGHTGWRVAI 743

QY 183 -----NOISGAHNTS-----DKQTPSKPKKQNSKGL-----APSVDFQOSVT----- 220
Db 744 SADGQRVSGSYDQTLRVWDAATGECVRELQGHSTSLVFAVALSADGQRIVSGSSDLTARV 803

QY 221 -----VLPQDENTLVSAVADGIIKWDLRKXNYTAYRQEPIASK 260
Db 804 WDTATGETLRELKGTGWRVSAFSTGQRIVTGGDDQSVRVWD-----AS- 849

QY 261 SFLYPGSSTRKL-GYS-----SLILDSTGTLFANCTDDNIYMFN-MTG-----LKTSPVA 309
Db 850 -----TGECVRELKGYTAALISVAFSPDQRIVSGGGDQTVRVVWNAATGECQCELK----- 900

QY 310 IFNGHONSTFYVKS-SLSPDDQFLVSGSSDEAAYIWKVSTPWQPPT-----VLLG 358
Db 901 ---GH---TEQVDSIAFSPDGHVSGSIDQTLRVWDVSSLSFSSSSSGGAGLRERQ 954

QY 359 HSGEVTVCWCPSPDFTKIATCSDNDTLKIWRNRG--LEKPGCDKLSTVGW 408
Db 955 HTKDVSNAF-PPDGKRLASGSDQSVRVWDVAVSGELLHELQHS-----GW 1000

Search completed: November 8, 2006, 18:19:38
Job time : 45 secs

This Page Blank (uspto/

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:14:54 ; Search time 183 Seconds
(without alignments)
1847.737 Million cell updates/sec

Title: US-10-726-160-2
Perfect score: 3877
Sequence: 1 MLFNSVLRQPQLGVLNRGWS.....TYFHRKSQDFCGPEHSTEL 730

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA Main:*
- 1: /EMC_Celerra_SID33/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /EMC_Celerra_SID33/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 3: /EMC_Celerra_SID33/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 4: /EMC_Celerra_SID33/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 5: /EMC_Celerra_SID33/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 6: /EMC_Celerra_SID33/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	730	4	US-10-145-396-12
2	3877	100.0	730	4	US-10-409-511-2
3	3877	100.0	730	5	US-10-726-160-2
4	3873	99.9	730	3	US-09-780-053-2
5	3503.5	90.4	673	5	US-10-330-773-674
6	3479.5	89.7	729	4	US-10-145-396-11
7	2914.5	75.2	784	5	US-10-330-773-671
8	1864	48.1	347	3	US-09-780-053-4
9	1165	30.0	522	4	US-10-403-571-74
10	1084	28.0	206	4	US-10-221-625-24
11	793.5	20.5	769	6	US-11-097-143-23652
12	759	19.6	146	5	US-10-330-773-676
13	753	19.4	351	3	US-09-780-053-5
14	587	15.1	523	5	US-10-739-930-8452
15	575.5	14.8	555	4	US-10-437-963-118806
16	460.5	11.9	314	4	US-10-424-599-282884
17	406	10.5	111	4	US-10-106-698-4768
18	312	8.0	610	4	US-10-369-493-20185
19	298	7.7	1005	4	US-10-369-493-20242
20	295.5	7.6	579	4	US-10-369-493-18893
21	281	7.2	1595	4	US-10-425-115-205747
22	279.5	7.2	1118	4	US-10-369-493-18845
23	269	6.9	598	4	US-10-425-115-221471
24	269	6.9	608	4	US-10-369-493-20224
25	267	6.9	1155	4	US-10-369-493-19869
26	266	6.9	606	4	US-10-369-493-18907
27	261	6.7	478	4	US-10-369-493-18905

28	259.5	6.7	1356	4	US-10-077-111-10	Sequence 10, Appl
29	258.5	6.7	1117	4	US-10-369-493-18965	Sequence 18965, A
30	256.5	6.6	343	4	US-10-119-932-5	Sequence 5, Appli
31	256	6.6	348	6	US-11-024-989-479	Sequence 479, App
32	256	6.6	515	4	US-10-369-493-1531	Sequence 1531, Ap
33	255.5	6.6	655	4	US-10-369-493-20066	Sequence 20066, A
34	254.5	6.6	1136	4	US-10-369-493-19046	Sequence 19046, A
35	254	6.6	379	4	US-10-369-493-8153	Sequence 8153, Ap
36	254	6.6	742	4	US-10-077-111-11	Sequence 11, Appl
37	253	6.5	375	4	US-10-119-932-1	Sequence 1, Appli
38	250	6.4	415	4	US-10-108-260A-4324	Sequence 4324, Ap
39	250	6.4	1140	4	US-10-369-493-19053	Sequence 19053, A
40	249.5	6.4	347	6	US-11-097-143-6498	Sequence 6498, Ap
41	249	6.4	329	4	US-10-424-599-272779	Sequence 272779, A
42	247.5	6.4	429	4	US-10-369-493-11348	Sequence 11348, A
43	247	6.4	559	4	US-10-369-493-20028	Sequence 20028, A
44	245	6.3	347	4	US-10-425-114-62553	Sequence 62553, A
45	244.5	6.3	626	4	US-10-369-493-20247	Sequence 20247, A

ALIGNMENTS

RESULT 1
US-10-145-396-12
; Sequence 12, Application US/10145396
; Publication No. US20030003485A1
; GENERAL INFORMATION:
; APPLICANT: Uenaka, Akiko
; TITLE OF INVENTION: METHODS FOR IDENTIFYING ANTIGENS
; FILE OF INVENTION: L00461/70129
; CURRENT APPLICATION NUMBER: US/10/145,396
; CURRENT FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 60/291,125
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-145-396-12

Query Match	100.0%;	Score 3877;	DB 4;	Length 730;
Best Local Similarity	100.0%;	Pred. No. 1.1e-259;	Mismatches 0;	Indels 0;
Matches 730;	Conservative 0;			Gaps 0;
Qy	1	MLFNSVLRQPQLGVLNRGWSQYPLQSLITGYQCSGNDEHTSYGETGVPPVPGCTFSSA	60	
Db	1	MLFNSVLRQPQLGVLNRGWSQYPLQSLITGYQCSGNDEHTSYGETGVPPVPGCTFSSA	60	
Qy	61	PNMEHLAVANBEGFVRLYNTESQSFRRKKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ	120	
Db	61	PNMEHLAVANBEGFVRLYNTESQSFRRKKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ	120	
Qy	121	TAKFVDVXAGELIGTCKGHQCSLKSVAFKPKVFCGGRDGNIMVMDTRCNKXKDDGFR	180	
Db	121	TAKFVDVXAGELIGTCKGHQCSLKSVAFKPKVFCGGRDGNIMVMDTRCNKXKDDGFR	180	
Qy	181	QVNQISGAHNTSDKOTPSKPKKKQNSKGLAPSVDFQQSVTVVLFDENTLVSAGAVDGI	240	
Db	181	QVNQISGAHNTSDKOTPSKPKKKQNSKGLAPSVDFQQSVTVVLFDENTLVSAGAVDGI	240	
Qy	241	KVWDLRKNYAYRQBPFIASKSLYPCGSSTRKLGYSLLDSTGFSTLFDNCTDNDIYMFNN	300	
Db	241	KVWDLRKNYAYRQBPFIASKSLYPCGSSTRKLGYSLLDSTGFSTLFDNCTDNDIYMFNN	300	
Qy	301	TGLKTSVPAIFNGHONSTFYVKSSLPDDQFLVSGSSDEAAIYKVPWPQPTVLLGHS	360	
Db	301	TGLKTSVPAIFNGHONSTFYVKSSLPDDQFLVSGSSDEAAIYKVPWPQPTVLLGHS	360	
Qy	361	QEVTSVCNCPDFTKIATCDDNTLKIWLNRGLBEPKGGDKLSTVGMASQKKESRPGL	420	

```
Db      361  QEVTSVCWCPDFTKIATCSDNTLKIWRLNRGLEKPGGDKLSTVGHASOKKESRGL 420
Qy      421  VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
Db      421  VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
Qy      481  NRRGSVSSVPKPPSPFPMKSIIRNWVTRTPSSPPPIPPASETKIMSPRKALIPVQKSSQ 540
Db      481  NRRGSVSSVPKPPSPFPMKSIIRNWVTRTPSSPPPIPPASETKIMSPRKALIPVQKSSQ 540
Qy      541  ABACSESRNRVRRRLDSSCLEVKQKVCVSCNCVTELDQOVENLHLDLCLAGNQEDLSK 600
Db      541  ABACSESRNRVRRRLDSSCLEVKQKVCVSCNCVTELDQOVENLHLDLCLAGNQEDLSK 600
Qy      601  DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLPCGEGSEMVGKENSSENK 660
Db      601  DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLPCGEGSEMVGKENSSENK 660
Qy      661  NWLLAMAAKRAENPSPSPSQTNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Db      661  NWLLAMAAKRAENPSPSPSQTNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Qy      721  FCGPEHSTEL 730
Db      721  FCGPEHSTEL 730
```

RESULT 2

```
US-10-409-511-2
; Sequence 2, Application US/10409511
; Publication No. US20030162268A1
; GENERAL INFORMATION:
; APPLICANT: IP, Nancy Y
; APPLICANT: CHEUNG, William M W
; TITLE OF INVENTION: No. US20030162268A1e1 Gene
; FILE REFERENCE: M98/0563/US
; CURRENT APPLICATION NUMBER: US/10/409,511
; CURRENT FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-409-511-2
```

```
Query Match      100.0%; Score 3877; DB 4; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.1e-259;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1  MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVPPVPPFGCTFSSA 60
Db      1  MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVPPVPPFGCTFSSA 60
Qy      61  PMEHVLAIVANEEGFRVRLYNTESQSPRKKCFKEWMAHNAVFDLAWVPGELKLVTAAGDQ 120
Db      61  PMEHVLAIVANEEGFRVRLYNTESQSPRKKCFKEWMAHNAVFDLAWVPGELKLVTAAGDQ 120
Qy      121  TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMWDTNCKNKGDFYR 180
Db      121  TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMWDTNCKNKGDFYR 180
Qy      181  QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQSVTVVLFQDENTLVSAGAVDGI 240
Db      181  QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQSVTVVLFQDENTLVSAGAVDGI 240
Qy      241  KWDLRKNTAYRQEPPIASKSFYPCGSTRKLGYSLLILDSTGSTLFPANCTDDNIYMFNM 300
Db      241  KWDLRKNTAYRQEPPIASKSFYPCGSTRKLGYSLLILDSTGSTLFPANCTDDNIYMFNM 300
Qy      301  TGLKTSPTVAIFNGHQNSTFYVKSSISPDQDFLVSGSSDEAAAYIWKVSTPWPQPTVLLGHS 360
```

```
Db      301  TGLKTSPTVAIFNGHQNSTFYVKSSISPDQDFLVSGSSDEAAAYIWKVSTPWPQPTVLLGHS 360
Qy      361  QEVTSVCWCPDFTKIATCSDNTLKIWRLNRGLEKPGGDKLSTVGHASOKKESRGL 420
Db      361  QEVTSVCWCPDFTKIATCSDNTLKIWRLNRGLEKPGGDKLSTVGHASOKKESRGL 420
Qy      421  VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
Db      421  VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
Qy      481  NRRGSVSSVPKPPSPFPMKSIIRNWVTRTPSSPPPIPPASETKIMSPRKALIPVQKSSQ 540
Db      481  NRRGSVSSVPKPPSPFPMKSIIRNWVTRTPSSPPPIPPASETKIMSPRKALIPVQKSSQ 540
Qy      541  ABACSESRNRVRRRLDSSCLEVKQKVCVSCNCVTELDQOVENLHLDLCLAGNQEDLSK 600
Db      541  ABACSESRNRVRRRLDSSCLEVKQKVCVSCNCVTELDQOVENLHLDLCLAGNQEDLSK 600
Qy      601  DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLPCGEGSEMVGKENSSENK 660
Db      601  DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLPCGEGSEMVGKENSSENK 660
Qy      661  NWLLAMAAKRAENPSPSPSQTNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Db      661  NWLLAMAAKRAENPSPSPSQTNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Qy      721  FCGPEHSTEL 730
Db      721  FCGPEHSTEL 730
```

RESULT 3

```
US-10-726-160-2
; Sequence 2, Application US/10726160
; Publication No. US20050037372A1
; GENERAL INFORMATION:
; APPLICANT: IP, Nancy Y
; APPLICANT: CHEUNG, William M W
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A RETINOIC ACID REGULATED PROTEIN
; FILE REFERENCE: FP4220B
; CURRENT APPLICATION NUMBER: US/10/726,160
; CURRENT FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: US 10/409,511
; PRIOR FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-726-160-2
```

```
Query Match      100.0%; Score 3877; DB 5; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.1e-259;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1  MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVPPVPPFGCTFSSA 60
Db      1  MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVPPVPPFGCTFSSA 60
Qy      61  PMEHVLAIVANEEGFRVRLYNTESQSPRKKCFKEWMAHNAVFDLAWVPGELKLVTAAGDQ 120
Db      61  PMEHVLAIVANEEGFRVRLYNTESQSPRKKCFKEWMAHNAVFDLAWVPGELKLVTAAGDQ 120
Qy      121  TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMWDTNCKNKGDFYR 180
Db      121  TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMWDTNCKNKGDFYR 180
Qy      181  QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQSVTVVLFQDENTLVSAGAVDGI 240
Db      181  QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQSVTVVLFQDENTLVSAGAVDGI 240
Qy      241  KWDLRKNTAYRQEPPIASKSFYPCGSTRKLGYSLLILDSTGSTLFPANCTDDNIYMFNM 300
```


Db 241 KWDLRKNTAYRQEPPIASKSFLYPGSTRKLGYSLLIDSTGSTLFANCTDDNIMFMN 300
Qy 301 TGLKTSFPAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
Db 301 TGLKTSFPAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
Qy 361 QEVTSVWCPCPSDFTKIATCSDNTLKIWLNRGLGEEKPGDKLSTVGMASQKKESRPG 420
Db 361 QEVTSVWCPCPSDFTKIATCSDNTLKIWLNRGLGEEKPGDKLSTVGMASQKKESRPG 420
Qy 421 VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPNTPTFSIKTSPAKARSPI 480
Db 421 VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPNTPTFSIKTSPAKARSPI 480
Qy 481 NREGSVSSVPKPPSFKMSIRNWTRTPSSSPPIPPASETKIMSPRKLIPVSKSSQ 540
Db 481 NREGSVSSVPKPPSFKMSIRNWTRTPSSSPPIPPASETKIMSPRKLIPVSKSSQ 540
Qy 541 AEACSESRNRVRRRLDSSCLESVKQKVCVSCNCTELDGOVENLHLDLCLAGNEDLSK 600
Db 541 AEACSESRNRVRRRLDSSCLESVKQKVCVSCNCTELDGOVENLHLDLCLAGNEDLSK 600
Qy 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGCGEGSEVMGKENSSENK 660
Db 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGCGEGSEVMGKENSSENK 660
Qy 661 NWLLAAAKRAENPRSPSSQTPNSRQSGKTLSPVTTITPSSMRKICTYFHRKSQED 720
Db 661 NWLLAAAKRAENPRSPSSQTPNSRQSGKTLSPVTTITPSSMRKICTYFHRKSQED 720
Qy 721 FCGPEHSTEL 730
Db 721 FCGPEHSTEL 730

RESULT 4

US-09-780-053-2
; Sequence 2, Application US/09780053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Paris
; APPLICANT: Eliana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Ava Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-780-053-2

Query Match 99.9%; Score 3873; DB 3; Length 730;
Best Local Similarity 99.9%; Pred. No. 2,1e-259;
Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MLFNSVLRQPOLGVLNRGWSQYPLQSLITGYQCSGNDEHTSYGETGVVPPPGCTFSSA 60
Db 1 MLFNSALRQPOLGVLNRGWSQYPLQSLITGYQCSGNDEHTSYGETGVVPPPGCTFSSA 60
Qy 61 PNMEHLAVANBEGFVRLYNTESQSRKCKFEMAHNAVFDLAWVPGLKLVTAAGDQ 120

Db 61 PNMEHLAVANBEGFVRLYNTESQSRKCKFEMAHNAVFDLAWVPGLKLVTAAGDQ 120
Qy 121 TAKFWDVWVAGELIGCTKGHCQSLKSAFSAKPKAVFCTGGRDGNIMVWDTRCNKKDGPYR 180
Db 121 TAKFWDVWVAGELIGCTKGHCQSLKSAFSAKPKAVFCTGGRDGNIMVWDTRCNKKDGPYR 180
Qy 181 QVNIQISGAHNTSDKQTPSKPKKKQNSKGLAPSVDFQQSVTVVLFDENTLVSAGAVDGI 240
Db 181 QVNIQISGAHNTSDKQTPSKPKKKQNSKGLAPSVDFQQSVTVVLFDENTLVSAGAVDGI 240
Qy 241 KWDLRKNTAYRQEPPIASKSFLYPGSTRKLGYSLLIDSTGSTLFANCTDDNIMFMN 300
Db 241 KWDLRKNTAYRQEPPIASKSFLYPGSTRKLGYSLLIDSTGSTLFANCTDDNIMFMN 300
Qy 301 TGLKTSFPAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
Db 301 TGLKTSFPAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
Qy 361 QEVTSVWCPCPSDFTKIATCSDNTLKIWLNRGLGEEKPGDKLSTVGMASQKKESRPG 420
Db 361 QEVTSVWCPCPSDFTKIATCSDNTLKIWLNRGLGEEKPGDKLSTVGMASQKKESRPG 420
Qy 421 VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPNTPTFSIKTSPAKARSPI 480
Db 421 VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPNTPTFSIKTSPAKARSPI 480
Qy 481 NREGSVSSVPKPPSFKMSIRNWTRTPSSSPPIPPASETKIMSPRKLIPVSKSSQ 540
Db 481 NREGSVSSVPKPPSFKMSIRNWTRTPSSSPPIPPASETKIMSPRKLIPVSKSSQ 540
Qy 541 AEACSESRNRVRRRLDSSCLESVKQKVCVSCNCTELDGOVENLHLDLCLAGNEDLSK 600
Db 541 AEACSESRNRVRRRLDSSCLESVKQKVCVSCNCTELDGOVENLHLDLCLAGNEDLSK 600
Qy 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGCGEGSEVMGKENSSENK 660
Db 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGCGEGSEVMGKENSSENK 660
Qy 661 NWLLAAAKRAENPRSPSSQTPNSRQSGKTLSPVTTITPSSMRKICTYFHRKSQED 720
Db 661 NWLLAAAKRAENPRSPSSQTPNSRQSGKTLSPVTTITPSSMRKICTYFHRKSQED 720
Qy 721 FCGPEHSTEL 730
Db 721 FCGPEHSTEL 730

RESULT 5

US-10-330-773-674
; Sequence 674, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 674
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-330-773-674

Query Match 90.4%; Score 3503.5; DB 5; Length 673;
Best Local Similarity 91.8%; Pred. No. 7.4e-234;
Matches 670; Conservative 1; Mismatches 2; Indels 57; Gaps 2;
Qy 1 MLFNSVLRQPOLGVLNRGWSQYPLQSLITGYQCSGNDEHTSYGETGVVPPPGCTFSSA 60
Db 1 MLFNSVLRQPOLGVLNRGWSQYPLQSLITGYQCSGNDEHTSYGETGVVPPPGCTFSSA 60

```
QY 61 PNMEHLAVANBEGFVRLYNTESQSRKCKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB 61 PNMEHLAVANBEGFVRLYNTESQSRKCKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
QY 121 TAKFWDVKAAGELIGTCKGHQCSLSKVAFSKFEKAVFCTGGRDGNIMVWDTNCKKDGFR 180
DB 121 TAKFWDVKAAGELIGTCKGHQCSLSKVAFSKFEK 156
QY 181 QVNOISGAHNTSDKOTPSKPKKQNSKGLAPSVDFQOSVTVVLFDQDENTLVSAGAVDGI 240
DB 157 -----DFQOSVTVVLFDQDENTLVSAGAVDGI 183
QY 241 KWDLRKNYTAHQEPIASKFELYPGSSTRKLGYSLLILDSTGTLFANCTDDNIYMFNM 300
DB 184 KWDLRKNYTAHQEPIASKFELYPGSSTRKLGYSLLILDSTGTLFANCTDDNIYMFNM 243
QY 301 TGLKTSFPAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB 244 TGLKTSFPAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 303
QY 361 QEVTSVCMCPSPDFTKIATCSDDNLTAKIWRNLNRLGLEEKPGDKLSTVGMASQKKESRPG 420
DB 304 QEVTSVCMCPSPDFTKIATCSDDNLTAKIWRNLNRLGLEEKPGDKLSTVGMASQKKESRPG 363
QY 421 VVTSSQSTPAKAPRVKCNPSNSPSSAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
DB 364 VVTSSQSTPAKAPRVKCNPSNSPSSAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 423
QY 481 NRRGVSVPKPPSPFKMSIRNWTPTSSPPITPPASETKIMSPKALIPVSKSSQ 540
DB 424 NRRGVSVPKPPSPFKMSIRNWTPTSSPPITPPASETKIMSPKALIPVSKSSQ 483
QY 541 AEACSESRNRVRRRLDSSCLSVKQCVKSCNCTELDQVENLHLDLCLAGNQEDLSK 600
DB 484 AEACSESRNRVRRRLDSSCLSVKQCVKSCNCTELDQVENLHLDLCLAGNQEDLSK 543
QY 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGEGSEVMVGKENSSENK 660
DB 544 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGEGSEVMVGKENSSENK 603
QY 661 NWLLAMAAKRAKNAENPSRPSQTPNSRRSQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
DB 604 NWLLAMAAKRAKNAENPSRPSQTPNSRRSQSGKTLPSVPTITPSSMRKICTYFHRKSQED 663
QY 721 FCGPEHSTEL 730
DB 664 FCGPEHSTEL 673
```

RESULT 6

```
US-10-145-396-11
; Sequence 11, Application US/10145396
; Publication No. US20030003485A1
; GENERAL INFORMATION:
; APPLICANT: Uenaka, Akiko
; APPLICANT: Nakayama, Eiichi
; TITLE OF INVENTION: METHODS FOR IDENTIFYING ANTIGENS
; FILE REFERENCE: L00461/70129
; CURRENT APPLICATION NUMBER: US/10/145,396
; CURRENT FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 60/291,125
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-145-396-11
```

Query Match 89.7%; Score 3479.5; DB 4; Length 729;
Best Local Similarity 89.2%; Pred. No. 3.8e-232;

```
Matches 651; Conservative 36; Mismatches 42; Indels 1; Gaps 1;
QY 1 MFNSVLRLQPOLGVLRNGWSSQYFQSLLTQYQCSGNDEHTSYGETGVVPVPPGCTFSSA 60
DB 1 MFNSVLRLQPOLGVLRNGWSSHYFQSLLSYQCNDEHTSYGETGVVPVPPGCTFCTA 60
QY 61 PNMEHLAVANBEGFVRLYNTESQSRKCKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB 61 PSMEHLAVANBEGFVRLYNTESQTSKKTCTCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
QY 121 TAKFWDVKAAGELIGTCKGHQCSLSKVAFSKFEKAVFCTGGRDGNIMVWDTNCKKDGFR 180
DB 121 TAKFWDVKAAGELIGTCKGHQCSLSKVAFPKQAVFSTGGRDGNIMIWDTRCNKKGFR 180
QY 181 QVNOISGAHNTSDKOTPSKPKKQNSKGLAPSVDFQOSVTVVLFDQDENTLVSAGAVDGI 240
DB 181 QVNOISGAHNTADKOTPSKPKKQNSKGLAPVDSQSVTVVLFDQDENTLVSAGAVDGI 240
QY 241 KWDLRKNYTAHQEPIASKFELYPGSSTRKLGYSLLILDSTGTLFANCTDDNIYMFNM 300
DB 241 KWDLRKNYTAHQEPIASKFELYPGSSTRKLGYSLLILDSTGTLFANCTDDNIYMENT 300
QY 301 TGLKTSFPAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB 301 TGLKTSFPAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSMHPHPPTVLLGHS 360
QY 361 QEVTSVCMCPSPDFTKIATCSDDNLTAKIWRNLNRLGLEEKPGDKLSTVGMASQKKESRPG 420
DB 361 QEVTSVCMCPSPDFTKIATCSDDNLTAKIWRNLNRLGLEEKPGDKLSTVGMASQKKESRPG 419
QY 421 VVTSSQSTPAKAPRVKCNPSNSPSSAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
DB 420 VVTSSQSTPAKAPRAKSSPSISPPSACTPSCAGDLPLPNTPTFSIKTSPATRSV 479
QY 481 NRRGVSVPKPPSPFKMSIRNWTPTSSPPITPPASETKIMSPKALIPVSKSSQ 540
DB 480 SRRGSISSVSPKPLSSFKMSIRNWTPTSSPPITPPASETKISSPKALIPVSKSSQ 539
QY 541 AEACSESRNRVRRRLDSSCLSVKQCVKSCNCTELDQVENLHLDLCLAGNQEDLSK 600
DB 540 ADACSESRNRVRRRLDSSCLSVKQCVKSCNCTELDQVABSLRLDCLSGTOEVL 599
QY 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGEGSEVMVGKENSSENK 660
DB 600 DSEGTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGEGSEVMVGKENSSENK 659
QY 661 NWLLAMAAKRAKNAENPSRPSQTPNSRRSQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
DB 660 NWLLAMAAKRAKNAENPSRPSQTPNSRRSQSGKTLPSVPTITPSSMRKICTYFHRKTQDD 719
QY 721 FCGPEHSTEL 730
DB 720 FCSPEHSTEL 729
```

RESULT 7

```
US-10-330-773-671
; Sequence 671, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 671
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-330-773-671
```

```

Query Match      75.2%; Score 2914.5; DB 5; Length 784;
Best Local Similarity 70.6%; Pred. No. 5.9e-193;
Matches 570; Conservative 34; Mismatches 46; Indels 157; Gaps 5;

QY 1 MLFNSVLRLQPOLGVLNR-----GNSQYPLQSLLTGYCQSGNDEH 40
DB 39 MLFNSVLRLQPOLGVLNRGAPPWIAADCGPGLPRPERIRASDRRTTSCLGSRFTLFTVI 98
QY 18 -----GNSQYPLQSLLTGYCQSGNDEH 40
DB 99 ELFSSRPKLGRSVASRAPGCLRQYRALTRCTTLVRVGGNSHYPLQSLLTGYCQSGNDEH 158
QY 41 TSGETGVPPVPPGCTFSSAPNMEHVLAVANEFGFVRLYNTESQSPKCKFKEMAHWNA 100
DB 159 TSGETGVPPVPPGCTFCTAPSMEHILAVANEFGFVRLYNTESQTSKTKCFK-----210
QY 101 VFDLAWVPGELKLVTAAGDQTAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFTCG 160
DB 211 -----GG 212
QY 161 RDGNIMWDTRCNKDGFRQVQVQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSVT 220
DB 213 RDGNIMWDTRCNKDGFRQVQVQISGAHNTADKQTPSKPKKQNSKGLAPVDSQOSVT 272
QY 221 VVLFQDENTLVSAAGAVDGIILKWDLRKNYTAAYRQEPITASKSFYLPGSSSTRKGLGYSSIL 280
DB 273 VVLFQDENTLVSAAGAVDGIILKWDLRKNYTAAYRQEPITASKSFYLPGSSSTRKGLGYSSIL 332
QY 281 STGSTLFPANCTDDNIYMFNMTGLKTSVPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSDEA 340
DB 333 STGSTLFPANCTDDNIYMFNMTGLKTSVPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSDEA 392
QY 341 AYIWKVSTPQOPPTVLLGHGSEVTSVWCPSDFTKIATCSDNTLKIWLNRGLGKPKGG 400
DB 393 AYIWKVSTPQOPPTVLLGHGSEVTSVWCPSDFTKIATCSDNTLKIWLNRGLGKPK-G 451
QY 401 DKLSTVCMASQKKESRPGILVTVTSOSTPAKAPRVKCNPSNPSAACAAPCAGDLPL 460
DB 452 DKHSIVGWTQKKREKACPVTVFSSQSTPAKAPRAKSPSISPSAACPSCAGDLPL 511
QY 461 PSNTPTFSIKTSPAKARSPINRRGSSVSPKPPSPFPMKIRNWTPTPSSPPITPPAS 520
DB 512 PSSTPTFSVKTPTATTRSSVSRGSSISSVSPKPLSLSPFMSLRNWTPTPSSPPITPPAS 571
QY 521 ETKIMSPRKALIPVQSKSQAECSESRNRVRRLLDSSCLSESVKQKCVKSCNCTELDQG 580
DB 572 ETKISSPRKALIPVQSKSQADACSESRNRVRRLLDSSCLSESVKQKCVKSCNCTELDQG 631
QY 581 VENLHLDLCLAGNOEDLSKDSLGPTKSKIEGAGTSISEPBPISPYASESGTLPPL 640
DB 632 AESRLDLCLLSGTQEVLSQDSSEGTSSKTEGAGTSISEPBPISPYASESGTLPPL 691
QY 641 RPCGEGSEMGKENSPPENKNWLLAIAAKKAENSPSPSPSSOTPNRRSGKTLSPVT 700
DB 692 RPCGEGSEMGKENSPPENKNWLLAIAAKKAENSPSPSPSSOTPNRRSGKTLSPVT 751
QY 701 ITPSSMRKIC-----TYFHRKSQEDFCG 723
DB 752 LA-----VCGNHQVLVDAGIKFCFCG 772

RESULT 8
US-09-780-053-4
; Sequence 4, Application US/09780053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Paris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell

```

```

; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83PSG4: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.SUSUI
; CURRENT APPLICATION NUMBER: US/09/780,053
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-780-053-4

Query Match      48.1%; Score 1864; DB 3; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.2e-120;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 PVPPGCTFSSAPNMEHVLAVANEFGFVRLYNTESQSPKCKFKEMAHWNAVFDLAWVP 108
DB 1 PVPPGCTFSSAPNMEHVLAVANEFGFVRLYNTESQSPKCKFKEMAHWNAVFDLAWVP 60
QY 109 GELKLVTAAGDQTAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFTCGRDNIMW 168
DB 61 GELKLVTAAGDQTAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFTCGRDNIMW 120
QY 169 DTRCNKKGDFYQVQVQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSVTVLFPDEN 228
DB 121 DTRCNKKGDFYQVQVQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSVTVLFPDEN 180
QY 229 TLVSAGAVDGIILKWDLRKNYTAAYRQEPITASKSFYLPGSSSTRKGLGYSSILSDSTGLFA 288
DB 181 TLVSAGAVDGIILKWDLRKNYTAAYRQEPITASKSFYLPGSSSTRKGLGYSSILSDSTGLFA 240
QY 289 NCTDDNIYMFNMTGLKTSVPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSDEAAYIWKVST 348
DB 241 NCTDDNIYMFNMTGLKTSVPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSDEAAYIWKVST 300
QY 349 PQOPPTVLLGHGSEVTSVWCPSDFTKIATCSDNTLKIWLNRGLG 395
DB 301 PQOPPTVLLGHGSEVTSVWCPSDFTKIATCSDNTLKIWLNRGLG 347

RESULT 9
US-10-403-571-74
; Sequence 74, Application US/10403571
; Publication No. US20040068763A1
; GENERAL INFORMATION:
; APPLICANT: Hopkins, Nancy
; APPLICANT: Golling, Gregory
; APPLICANT: Amsterdam, Adam
; APPLICANT: Sun, Zhaoxia
; TITLE OF INVENTION: Developmental Mutations in Zebrafish
; FILE REFERENCE: 01997/539002
; CURRENT APPLICATION NUMBER: US/10/403,571
; CURRENT FILING DATE: 2003-03-25
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; TYPE: PRT
; ORGANISM: Danio rerio
; US-10-403-571-74

Query Match      30.0%; Score 1165; DB 4; Length 522;
Best Local Similarity 38.1%; Pred. No. 5.5e-72;
Matches 275; Conservative 74; Mismatches 139; Indels 234; Gaps 20;

QY 16 RNCWSSQYPLQSLLTGYCQSGNDEHTSYGTVPPFPFGCTFSSAPNMEHVLAVANEBOG 75

```

Db	17	RNGEORLFLPLSLLDCEYECARRDEHISYGASAAAVPPFGCTSSAHQQOCLAVANEFG	76
Qy	76	VRLYNT-ESQSPRKKCFKEWAHNAVFDLAWPGBELKVATAAGDOTAKFMDVKACELIG	134
Db	77	VTFINTEGKQS---SVLKEMQAHNDNAVEDIAWVPTNCLVTASGDQFARLMDVITGDLLG	133
Qy	135	TCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTRCNKDGFYRNOVIGSAHNSTDK	194
Db	134	TFKGHCLSKSVAFYKQEKAVFSTGGRDGNIMIWDTRCSKKDGFYRQVKVIGSAHMKPER	193
Qy	195	QTPSKPKKKQNSKGLAPSVDFQOSVTVVLFDQENTLVSGAVDGIKVMDLRKNYTAYRQ	254
Db	194	FTPQTKR---RGMAPPVDSQGVTVVLFCDETKLSSGAVDGIKVMDLRKNYTAYHQ	249
Qy	255	EPIASKSPLYPGSSTRKLGYSLLDSTGSTLPANCTDDNIYNFNTMTGLKTSVPAIFNGH	314
Db	250	NPLPLQAYPGSCTRKLGYSGLSLDYTGSRLFNSCTDDNIYNFNISGLKTTPIAVFSGH	309
Qy	315	QNSTFVVKSSLPDDQFLVSGSSDEAAYIKVSVTPWQPPVTVLGHSEVTSVCMCPSDT	374
Db	310	SNSSFYKSTVSPD-----SYRCRFTPP-----IKH-----LFCFS-----	340
Qy	375	KIATCSDDDNTLKIWRLNRGLEBKPGDGKLVSTVGASQKKESRPLGTVTSSOSTPAKAP	434
Db	341	-----SQAHQFONAIIPSTWDPQOQ---360	
Qy	435	RVKCNPSNSPSAACAPCSAGDLPLPNTPTFTSIKTSPAKARSPINRRGSVSSVPKPP	494
Db	361	QISC-----SKSPHTCSP-----GS-----375	
Qy	495	SSPQMSIRWVTRTPSSSPPTTPASETKIMSPRKALIPVQKSQAACESESRNVKR	554
Db	376	-----VLRTPCQTP-----TGDRFCQLWFGRGDR	401
Qy	555	LDSSCLLESYKQKVCVSCNVTELDQOVENHLDLCLLAGNOEDLSKDSLGPTKSSKIEGA	614
Db	402	-----CVLYPN-VKRSRGSVTLK-----KEDSFG-LESERKQGS	434
Qy	615	GTSISEPPSPISPYASESCGTLPLRPCGBGSEVMGKENSPPENKNWLLAMAARAKAN	674
Db	435	-----DGAASGKENSPPRTDWLSVQKFKG-S	463
Qy	675	PSRPSSTQTPNSRRQSGKTLPSVTTTP-----SSMKKICTYFHRKS	717
Db	464	AQPKSFSS--GSSQDTRTLESAAVSPRPMKVPSPPTNKASPSKPMKISSYPMKRT	520
Qy	718	QE 719	
Db	521	QD 522	

```

RESULT 10
US-10-221-025-24
Sequence 24, Application US/10221162S
Publication No. US20040033942A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: YUE, Henry
APPLICANT: LAL, Preeti
APPLICANT: LU, Dyung Aina M.
APPLICANT: PATTTERSON, Chandra
APPLICANT: AZIMZAI, Yalda
APPLICANT: BANDMAN, Olga
APPLICANT: TANG, Y. Tom
APPLICANT: MATHUR, Preete
APPLICANT: SHAH, Purvi
APPLICANT: AU-YOUNG, Janice
APPLICANT: REDDY, Roopa
TITLE OF INVENTION: TRANSCRIPTION F
FILE REFERENCE: PF-0761 PCT

```

```

; CURRENT APPLICATION NUMBER: US/10/221,625
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040033942A1 1377380CD1
US-10-221-625-24

Query Match      28.0%; Score 1084; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 6.4e-67;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps

Qy      525  MSPRKALIPVSKSQAEACSESRNRVKKRLDSSCLSEVKQKCVKSCNCVTELDGQVNL
Db      1   MSPRKALIPVSKSQAEACSESRNRVKKRLDSSCLSEVKQKCVKSCNCVTELDGQVNL
Qy      585  HLDLCLAGNQEDLSKDSIGPTKSSKIEGAGTISEPPSPISPYASESGTLPPLPFCG
Db      61  HLDLCLAGNQEDLSKDSIGPTKSSKIEGAGTISEPPSPISPYASESGTLPPLPFCG
Qy      645  EGSEWVGKNSPENGKWLWLAAMAKRKAENPSPSPSQTPNRRRQSGKTLPSPTVTTPS
Db      121  EGSEWVGKNSPENGKWLWLAAMAKRKAENPSPSPSQTPNRRRQSGKTLPSPTVTTPS
Qy      705  SMRKICTYFHRKSOEDFCGPEHSTEL 730
Db      181  SMRKICTYFHRKSOEDFCGPEHSTEL 206

RESULT 11
US-11-097-143-23652
; Sequence 23652, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.

```

```

RESULT 11
US-11-097-143-23652
; Sequence 23652, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23652
; LENGTH: 769
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-23652

Query Match          20.5%; Score 793.5; DB 6; Length 769;
Best Local Similarity 29.3%; Pred. No. 5.2e-46;

```

```

Matches 224; Conservative 112; Mismatches 268; Indels 163; Gaps 25;
QY 49 PVPP-FGCTFSSAPNMEHVLAVANESGFVRLYNT-----ESQSF-RKKCFKEMAHWN 99
Db 49 PEPIPSAKFANCDGVRHILAIANEDGKITLQDTTQRNHQPEEQSLVGPQC-----HEN 102
QY 100 AVFDLAWVPGEKLVTAAGDQTAFAFDVWKAGELIG--TCKGHQCSLKSVAFSKFEKAVFC 157
Db 103 AVFDLEWAPQGRFVSASGDHTARLWEVAGSGIRGLNSYVGHTRSVKSAAFKRTDPAVFA 162
QY 158 TGRDGNIMVWDRCTNCKKQDFYRVQNOISGAHNTSDKQTPSKPKKONSKGLAPSVDFOQ 217
Db 163 TGRDGAIIWDIRANLNMDLTSRVNCIYSGHTGPGTVPVQKQRTTRPKMAGTTSS 222
QY 218 SVTVLVFQDENTLVSAGAVDGIKIVMDLRKNYTAHQEPIASKSFYPGSSTRKLGYSYL 277
Db 223 SITGLAFQNDTLISCGAGDGVIKVWDLRRNYTAYKKEPLPRHKLPRYAGSSTFR-GFTNL 281
QY 278 ILDTSTGTLFANCTDDNIYFMNMTGLKTSFVAIFNGHONSTFYVKSLSLSPDDQFLVSGSS 337
Db 282 IVDASGTRLYANCMDNTIYCYNLASYSQRLACYKGLNSTFYIKSLSPDQKGLLSGSS 341
QY 338 DEAYITWVSTPMQPTVLGHQSOVTSVCMCPDFTKIATCDDNTLKIWRNL-----391
Db 342 DERAYIWNLDHAEPLVALAGHTVEVTCVAMGSSHDCPIVTCDDARHKIWRIGPDLGL 401
QY 392 -----RGLBEKPGDK-----402
Db 402 SEAEBAEKYRGTAHYREFGKKAFGSPSGNHNKLNRLDLESTPRSLKRLMDQNERTPGSVE 461
QY 403 -----LSTVGWASQ-----KKESRPLGLVTVTSSQSTPAKAPVK-CNPS 441
Db 462 KITTKSFLEMLGVAQTEATEQOKRAKPLESRGRRLFGPSSQETACHEHQLQLOINEE 521
QY 442 NSPSSAACAPSCAGDLPLPSNTPTTSIKTSPAKARSPINRRGSSVSPKPPSPKMSI 501
Db 522 DASPSKRQKNSAAEDV-----SPLHKLSTP--SHSPLSE--NVNHYVTSPTPTS----568
QY 502 RNWVTTPSSPPIPTPASETKIMSPRKALI PVSQSSQAE--ACSESRRVKKRLDSSC 559
Db 569 -----AAAAVAADALNPPISAAIYSPTSNL-PNYVLDEGAPHLGIMSPKRAKEKYD--W 622
QY 560 LESVKOKVKVSCNCVTELDQVENHLDLCLAGNOEDLSKDSGLGPTKSSKIEGAGTSI 619
Db 623 LTNIRKQKLSGRAHTLSEKI-----SEEQADVLA---SPRLQSLRQSEC 666
QY 620 EPPSPISP-----YASESCGTLPLPLRPGCGSEMVGKENSPPENKNWLLAMAARKAEN 674
Db 667 SPRIHASPRRRIHTDGGGT-----PAGSSSH--SHSQSQPKTPT-----SSRNSSET 713
QY 675 P-----SPRSPSSQTPNSRRQSGKTLPSVPTI-----TPSSMR 707
Db 714 TLLRFFSIQRSSVPAEETTTTNAAPSSSDPHPAVTAAPATPLSMR 760

RESULT 12
US-10-330-773-676
; Sequence 676, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 676
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: VARIANT
; LOCATION: (1)....(146)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-330-773-676

Query Match 19.6%; Score 759; DB 5; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 499 MSIRWVTRTPSSSPITPPASETKIMSPRKALIPVSQKSSQAEACSESRNVKRLDSS 558
Db 1 MSIRWVTRTPSSSPITPPASETKIMSPRKALIPVSQKSSQAEACSESRNVKRLDSS 60
QY 559 CUESVKOKVKVSCNCVTELDQVENHLDLCLAGNOEDLSKDSGLGPTKSSKIEGAGTSI 618
Db 61 CUESVKOKVKVSCNCVTELDQVENHLDLCLAGNOEDLSKDSGLGPTKSSKIEGAGTSI 120
QY 619 SEPPSPISPYASESCGTLPLPLRPC 643
Db 121 SEPPSPISPYASESCGTLPLPLRPC 145

RESULT 13
US-09-780-053-5
; Sequence 5, Application US/09780053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Bid
; APPLICANT: Mary Faris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-780-053-5

Query Match 19.4%; Score 753; DB 3; Length 351;
Best Local Similarity 42.7%; Pred. No. 1.1e-43;
Matches 153; Conservative 62; Mismatches 125; Indels 18; Gaps 6;

QY 49 PVPP-FGCTFSSAPNMEHVLAVANESGFVRLYNT-----ESQSF-RKKCFKEMAHWN 99
Db 1 PEPIPSAKFANCDGVRHILAIANEDGKITLQDTTQRNHQPEEQSLVGPQC-----HEN 54
QY 100 AVFDLAWVPGEKLVTAAGDQTAFAFDVWKAGELIG--TCKGHQCSLKSVAFSKFEKAVFC 157
Db 55 AVFDLEWAPQGRFVSASGDHTARLWEVAGSGIRGLNSYVGHTRSVKSAAFKRTDPAVFA 114
QY 158 TGRDGNIMVWDRCTNCKKQDFYRVQNOISGAHNTSDKQTPSKPKKONSKGLAPSVDFOQ 217
Db 115 TGRDGAIIWDIRANLNMDLTSRVNCIYSGHTGPGTVPVQKQRTTRPKMAGTTSS 174
QY 218 SVTVLVFQDENTLVSAGAVDGIKIVMDLRKNYTAHQEPIASKSFYPGSSTRKLGYSYL 277
Db 175 SITGLAFQNDTLISCGAGDGVIKVWDLRRNYTAYKKEPLPRHKLPRYAGSSTFR-GFTNL 233
QY 278 ILDTSTGTLFANCTDDNIYFMNMTGLKTSFVAIFNGHONSTFYVKSLSLSPDDQFLVSGSS 337
Db 234 IVDASGTRLYANCMDNTIYCYNLASYSRPRACYKGLNSTFYIKSLSPDQKGLLSGSS 293

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:14:04 ; Search time 53 Seconds
(without alignments)
1205.611 Million cell updates/sec

Title: US-10-726-160-2

Perfect score: 3877

Sequence: 1 MLFNSVLRPQLGLVRLNGWS.....TYFHRKSQEDFCGPEHSTEL 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters:: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /EMC Celerra_SIDS3/ptodata/2/iaa/5 COMB.pcp:*
2: /EMC Celerra_SIDS3/ptodata/2/iaa/6 COMB.pcp:*
3: /EMC Celerra_SIDS3/ptodata/2/iaa/7 COMB.pcp:*
4: /EMC Celerra_SIDS3/ptodata/2/iaa/H COMB.pcp:*
5: /EMC Celerra_SIDS3/ptodata/2/iaa/PCTUS COMB.pcp:*
6: /EMC Celerra_SIDS3/ptodata/2/iaa/RE COMB.pcp:*
7: /EMC Celerra_SIDS3/ptodata/2/iaa/baGfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256.5	6.6	343	2	US-09-063-743-5
2	256.5	6.6	343	2	US-09-590-540-5
3	253	6.5	375	2	US-09-063-743-1
4	253	6.5	375	2	US-09-590-540-1
5	245.5	6.3	514	1	US-08-190-802A-66
6	245.5	6.3	514	2	US-08-477-346-66
7	245.5	6.3	514	2	US-08-473-089-66
8	245.5	6.3	514	2	US-08-487-072A-66
9	241	6.2	251	2	US-09-291-170A-13
10	241	6.2	251	2	US-09-724-884-13
11	241	6.2	251	2	US-09-724-592-13
12	241	6.2	251	2	US-09-673-222-13
13	233.5	6.0	606	1	US-08-883-534-3
14	233.5	6.0	606	2	US-09-204-764-3
15	233.5	6.0	627	2	US-09-949-016-7082
16	233.5	6.0	627	2	US-09-949-016-7083
17	230.5	5.9	416	2	US-09-270-767-41646
18	211.5	5.5	587	2	US-08-899-578-2
19	209.5	5.4	395	2	US-09-032-372-1
20	208.5	5.4	409	1	US-08-283-917-3
21	208.5	5.4	409	1	US-08-861-716-3
22	208.5	5.4	409	2	US-09-538-092-1119
23	208.5	5.4	410	1	US-08-283-917-9
24	208.5	5.4	410	1	US-08-961-716-9
25	207.5	5.4	690	2	US-09-724-592-2
26	207.5	5.4	690	2	US-09-673-222-2

27	207	5.3	423	2	US-09-013-118-1	Sequence 1, Appli
28	206.5	5.3	690	2	US-09-291-170A-2	Sequence 2, Appli
29	206.5	5.3	690	2	US-09-724-884-2	Sequence 2, Appli
30	201.5	5.2	540	2	US-09-213-888-7	Sequence 7, Appli
31	201.5	5.2	540	2	US-09-213-888-10	Sequence 10, Appli
32	201.5	5.2	540	2	US-09-328-877D-7	Sequence 7, Appli
33	201.5	5.2	540	2	US-09-328-877D-10	Sequence 10, Appli
34	201.5	5.2	545	2	US-09-213-888-6	Sequence 6, Appli
35	201.5	5.2	545	2	US-09-328-877D-6	Sequence 6, Appli
36	201.5	5.2	553	2	US-09-213-888-5	Sequence 5, Appli
37	201.5	5.2	553	2	US-09-328-877D-5	Sequence 5, Appli
38	201.5	5.2	559	2	US-09-213-888-9	Sequence 9, Appli
39	201.5	5.2	559	2	US-09-328-877D-9	Sequence 9, Appli
40	201.5	5.2	589	2	US-09-213-888-8	Sequence 8, Appli
41	201.5	5.2	589	2	US-09-328-877D-8	Sequence 8, Appli
42	201.5	5.2	592	2	US-09-213-888-4	Sequence 4, Appli
43	201.5	5.2	592	2	US-09-328-877D-4	Sequence 4, Appli
44	201.5	5.2	626	2	US-09-213-888-21	Sequence 21, Appli
45	201.5	5.2	626	2	US-09-328-877D-21	Sequence 21, Appli

ALIGNMENTS

RESULT 1
US-09-063-743-5
; Sequence 5, Application US/09063743
; Patent No. 6242214
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/063,743
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Carrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0508 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 2281093
US-09-063-743-5

```
Query Match 6.6%; Score 256.5; DB 2; Length 343;
Best Local Similarity 25.6%; Pred. No. 1.9e-12;
Matches 79; Conservative 44; Mismatches 117; Indels 69; Gaps 9;

QY 89 KCFKEMAHNAVFDLAWVPGELKLVTAAGDQDTAKFWDVKAGELIGTCKGHQCSLKSVAF 148
DB 87 KNFMVLKGHNAILDLHWTSDGSIIVSAPDKTVRAWDVETGKQIKMAEHSFVNSCCP 146
QY 149 SKFEKAVFCTGGRGDGNIMVWDTKCNKKGDFYRQVNIQISGAHNTSDKQTPSKPKKQNSKG 208
DB 147 TRRGPPLIISGDDGTAKLWDMR-----QRG 172
QY 209 LAPSVDFQSVTVVLVFDQENTLVLSAGAVDGIKVDLKNYTYAYRQEPISKSLFLPGSS 268
DB 173 AIQTTPDKYQITAVSFSDAADKIFTGGVDNDVKVWDLRK-----GEA 214
QY 269 TRKL-GYSSLI-----LDSTGSTLPANCTDDNIYMFNMT--GLKTSVPAIFNGHONS--TF 319
DB 215 TMTLEGHQDTITGMSLSPDGSYLLTNGMDNKLVCVDMRPYAPQNCVKIFEGHQHNFKN 274
QY 320 YKSSLSPPDQFLVSGSDEAAYIWKVS---TPWQPPTVLLGHQSQEVTSVCWCPSDFTKI 376
DB 275 LLKCSWSPDGTKVYTAGSSDRWVHIWDTTSRTIYKLP-----GHTGSVNECVFHPTE-PII 329
QY 377 ATCSDNTL 385
DB 330 GSCSSDKNI 338

RESULT 2
US-09-590-540-5
; Sequence 5, Application US/09590540
; Patent No. 6410267
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Lal, Preeti
; Guegler, Karl J.
; Corley, Neil C.
; Patterson, Chandra
; TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Genomics, Inc.
; STREET: 3160 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/590,540
; FILING DATE: 08-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/063,743
; FILING DATE: April 21, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: David G. Streeter
; REGISTRATION NUMBER: 43,168
; REFERENCE/DOCKET NUMBER: PF-0508-1 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
;
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 2281093
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-590-540-5

Query Match 6.6%; Score 256.5; DB 2; Length 343;
Best Local Similarity 25.6%; Pred. No. 1.9e-12;
Matches 79; Conservative 44; Mismatches 117; Indels 69; Gaps 9;

QY 89 KCFKEMAHNAVFDLAWVPGELKLVTAAGDQDTAKFWDVKAGELIGTCKGHQCSLKSVAF 148
DB 87 KNFMVLKGHNAILDLHWTSDGSIIVSAPDKTVRAWDVETGKQIKMAEHSFVNSCCP 146
QY 149 SKFEKAVFCTGGRGDGNIMVWDTKCNKKGDFYRQVNIQISGAHNTSDKQTPSKPKKQNSKG 208
DB 147 TRRGPPLIISGDDGTAKLWDMR-----QRG 172
QY 209 LAPSVDFQSVTVVLVFDQENTLVLSAGAVDGIKVDLKNYTYAYRQEPISKSLFLPGSS 268
DB 173 AIQTTPDKYQITAVSFSDAADKIFTGGVDNDVKVWDLRK-----GEA 214
QY 269 TRKL-GYSSLI-----LDSTGSTLPANCTDDNIYMFNMT--GLKTSVPAIFNGHONS--TF 319
DB 215 TMTLEGHQDTITGMSLSPDGSYLLTNGMDNKLVCVDMRPYAPQNCVKIFEGHQHNFKN 274
QY 320 YKSSLSPPDQFLVSGSDEAAYIWKVS---TPWQPPTVLLGHQSQEVTSVCWCPSDFTKI 376
DB 275 LLKCSWSPDGTKVYTAGSSDRWVHIWDTTSRTIYKLP-----GHTGSVNECVFHPTE-PII 329
QY 377 ATCSDNTL 385
DB 330 GSCSSDKNI 338

RESULT 3
US-09-063-743-1
; Sequence 1, Application US/09063743
; Patent No. 6242214
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/063,743
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0508 US
; TELECOMMUNICATION INFORMATION:
```

```
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGNOT01
CLONE: 059953
US-09-063-743-1

Query Match
Best Local Similarity 25.7%; Score 253; DB 2; Length 375;
Matches 98; Conservative 49; Mismatches 159; Indels 76; Gaps 14;

QY 20 SSQYPLQSLLTGYQCSGNDEHTSYGETGVPPFGCTFSSAPNMEHVLAVANEEGFVRLY 79
DB 54 SLQAPIM-LLSGHE-----GEV-----YCKEFH--PN-GSTLASAGFDRLLILW 93
QY 80 NTESQSRKKCFKEWMAHNAVFDLAWVPEGLKLVTAAGDQTAKEFWDVKAGELIGTCCKGH 139
DB 94 NVYGCDCNYATLK---GHSGAVMELHNTDGSMLFSASTDKTVAVWVDSGTSERVKRLKGH 150
QY 140 QCSLKSVAPEKFAVCTGGRDGNIMVWDTRCNKKGDFYRQVNOISGAHNTSDKQTPSK 199
DB 151 TSFVNSCYPARRGQPVLCVTSDDGTGKWDIR--KKA-----IQTQNT----- 193
QY 200 PKKQNSKGLAPSVDFQOSVTVVLFQDENTLVSAAGVDGIIKVDLRKNVYAYRQSPIAS 259
DB 194 -----YQVLAVTFNDTSDQIIISGIDNDIKVDLRQNKLYTMRGHAD 236
QY 260 KSFLYPGSSTRKLGYSLLDSTGLFANCTDDNIYFNMTGL--KTSFVAIFNG--HQ 315
DB 237 -----SVTGLSLSSGYSLLSNAMNTVRVMDVRFPAPKRCVKIFQGNVHN 283
QY 316 NSTFYVYKSSLPDQPLVSGSSDEAAYIKWVSTPWPQPTVLLGHSEVTSVCWCPSDFTK 375
DB 284 FEKNLRCWSVDPGSKIAGSADRFVYVWD--TTSRRILYKLPGHAGSINEVAFHP--DEPI 341
QY 376 IATCSDNTLKIWLNRGLEEK 397
DB 342 IISASDKKLYMGFEFSEDMDWK 363

RESULT 4
US-09-590-540-1
; Sequence 1, Application US/09590540
; Patent No. 6410267
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Lali, Preeti
; Guegler, Karl J.
; Corley, Neil C.
; Patterson, Chandra
; TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Genomics, Inc.
; STREET: 1160 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/590,540

TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGNOT01
CLONE: 059953
US-09-063-743-1

Query Match
Best Local Similarity 25.7%; Score 253; DB 2; Length 375;
Matches 98; Conservative 49; Mismatches 159; Indels 76; Gaps 14;

QY 20 SSQYPLQSLLTGYQCSGNDEHTSYGETGVPPFGCTFSSAPNMEHVLAVANEEGFVRLY 79
DB 54 SLQAPIM-LLSGHE-----GEV-----YCKEFH--PN-GSTLASAGFDRLLILW 93
QY 80 NTESQSRKKCFKEWMAHNAVFDLAWVPEGLKLVTAAGDQTAKEFWDVKAGELIGTCCKGH 139
DB 94 NVYGCDCNYATLK---GHSGAVMELHNTDGSMLFSASTDKTVAVWVDSGTSERVKRLKGH 150
QY 140 QCSLKSVAPEKFAVCTGGRDGNIMVWDTRCNKKGDFYRQVNOISGAHNTSDKQTPSK 199
DB 151 TSFVNSCYPARRGQPVLCVTSDDGTGKWDIR--KKA-----IQTQNT----- 193
QY 200 PKKQNSKGLAPSVDFQOSVTVVLFQDENTLVSAAGVDGIIKVDLRKNVYAYRQSPIAS 259
DB 194 -----YQVLAVTFNDTSDQIIISGIDNDIKVDLRQNKLYTMRGHAD 236
QY 260 KSFLYPGSSTRKLGYSLLDSTGLFANCTDDNIYFNMTGL--KTSFVAIFNG--HQ 315
DB 237 -----SVTGLSLSSGYSLLSNAMNTVRVMDVRFPAPKRCVKIFQGNVHN 283
QY 316 NSTFYVYKSSLPDQPLVSGSSDEAAYIKWVSTPWPQPTVLLGHSEVTSVCWCPSDFTK 375
DB 284 FEKNLRCWSVDPGSKIAGSADRFVYVWD--TTSRRILYKLPGHAGSINEVAFHP--DEPI 341
QY 376 IATCSDNTLKIWLNRGLEEK 397
DB 342 IISASDKKLYMGFEFSEDMDWK 363

RESULT 5
US-08-190-802A-66
; Sequence 66, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
```

```
/ COUNTRY: USA
/ ZIP: 94306-0850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ FILING DATE: 01-FEB-1994
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 8600-0139
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 66:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 514 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49
/ US-08-190-802A-66

Query Match
Best Local Similarity 24.1%; Score 245.5; DB 1; Length 514;
Matches 116; Conservative 62; Mismatches 196; Indels 107; Gaps 23;

QY 13 GVLR-NGWSSQYPLQSLTGYQCSGNDEHTSYGETGVVP-PFGCTFFSAPNMEHVLAV- 69
DB 47 GAIURVGAISEKQLELLN--QLNGTSDD-----PVPYTFSCITQKKASDPVKITD 96
QY 70 -----ANEGFVRLNNTESQSPRKKCFKEMMA---HWNAPVFLAWVP-GE 110
DB 97 ITDNLVSSLLKPGYNSTEDQITLLYTPRAVFKVPVTRSSSAIAGHGSTILCSAFAPHTS 156
QY 111 LKLVTAAGDQTAQFWDVVRAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDT 170
DB 157 SRMVTGAGDNTARINWCDTQTPMHTLKGHYNWVLCVSWSP-DGEVIATGSDMTIRLWDP 215
QY 171 RCKKKDGYRQVNOISGAHN---TSDKQTP-----SKPKKKQNSKGLAPSV----- 213
DB 216 KSGQCLG-----DALRG-HSKWITSLSWEPILHVPKSPKRLASSSSKDGTIKIWDTVSRV 269
QY 214 -----DFQOSVTVVLFQDENTLVSAAGVDGIKVKWDLR-----KNYTAIRQEPIA 258
DB 329 STDYALRIGAFDHTGKTPSTPEEAQKALENYEKICKKNGNSEEMVMTASDDYTWFLWNP 388
QY 303 LK-TSPVATFNHONSTFYVKSLSLDDQFLVSGSSDEAAYTWK-----VSTPWQPPTV 355
DB 389 LKSTKPIARMTGHOKLVNHV--AFSPDGRYIVYASFDNSIKLWDGRDGRGKFIST----- 439
QY 356 LLGHSEVTSVCMPCSPDFTKIATCSDDNTLKIW--RLNRGLBEKPGCDKSLSTVGWASQKK 413
DB 440 FRGHIAVQVAV--SSDCRLVSCSKDTTLKVDVTRKLSVDLPQGIKTKLYVDNSVDGK 498
QY 414 K 414
DB 499 R 499

RESULT 6
US-08-477-346-66
```

```
/ Sequence 66, Application US/08477346
/ Patent No. 6262023
/ GENERAL INFORMATION:
/ APPLICANT: Mochly-Rosen, Daria
/ APPLICANT: Ron, Dorit
/ TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
/ TITLE OF INVENTION: Thereof
/ NUMBER OF SEQUENCES: 265
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Morrison & Foerster
/ STREET: 2000 Pennsylvania Avenue, NW
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20006-1812
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/477,346
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/487,072
/ FILING DATE: 07-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MURASHIGE, KATE H.
/ REGISTRATION NUMBER: 29,959
/ REFERENCE/DOCKET NUMBER: 2550-0025.20
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 887-1500
/ TELEFAX: (202) 887-0763
/ INFORMATION FOR SEQ ID NO: 66:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 514 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49
/ US-08-477-346-66
```

```
Query Match
Best Local Similarity 24.1%; Score 245.5; DB 2; Length 514;
Matches 116; Conservative 62; Mismatches 196; Indels 107; Gaps 23;

QY 13 GVLR-NGWSSQYPLQSLTGYQCSGNDEHTSYGETGVVP-PFGCTFFSAPNMEHVLAV- 69
DB 47 GAIURVGAISEKQLELLN--QLNGTSDD-----PVPYTFSCITQKKASDPVKITD 96
QY 70 -----ANEGFVRLNNTESQSPRKKCFKEMMA---HWNAPVFLAWVP-GE 110
DB 97 ITDNLVSSLLKPGYNSTEDQITLLYTPRAVFKVPVTRSSSAIAGHGSTILCSAFAPHTS 156
QY 111 LKLVTAAGDQTAQFWDVVRAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDT 170
DB 157 SRMVTGAGDNTARINWCDTQTPMHTLKGHYNWVLCVSWSP-DGEVIATGSDMTIRLWDP 215
QY 171 RCKKKDGYRQVNOISGAHN---TSDKQTP-----SKPKKKQNSKGLAPSV----- 213
DB 216 KSGQCLG-----DALRG-HSKWITSLSWEPILHVPKSPKRLASSSSKDGTIKIWDTVSRV 269
QY 214 -----DFQOSVTVVLFQDENTLVSAAGVDGIKVKWDLR-----KNYTAIRQEPIA 258
DB 270 COYTMSGHTNSVSCVKWGGQGLLYS-GSHDRTVRVWDINSQGRNCINILKSHAHWNHLSL 328
QY 259 SKSFLY-----PGSSSTRKL--GYSSLLDSTGTSTLFCNCTDDNNIYFMNMTG 302
DB 329 STDYALRIGAFDHTGKTPSTPEEAQKALENYEKICKKNGNSEEMVMTASDDYTWFLWNP 388
```



```

Best Local Similarity   24.1%; Pred. No. 2.8e-11;
Matches 116; Conservative 62; Mismatches 196; Indels 107; Gaps 23;

QY      13 GVLR-NGWSSVPIQLSLLTGQCNGNDHTSYGETGVDPV-PFGCTFSSAPNMEHVLAV- 69
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       47 GALRVPGAISEKQLEELLN--QLNGTSD-----PVPYTFSCITGCKKASBPVKTIID 96

QY      70 -----ANEEGFVRLYNTESQSFRKKCFKEMMA---HMNAVFDLAWVP-GE 110
         :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db       97 ITDNLYSLLIKPGYNSTEDQITLLYTTPRAVFKVPVTRSSSAIAGHGSTITCSAFAPHTS 156

QY     111 LKLVTAAGDQTAKFWDVVAGELIGTCKGHCQSLKSVAFSKFKEKAVFCTGGRDGNIMWDT 170
         :||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     157 SRMVTGAGDNTARIDWCDDTOTPMHTLKGHYNVLCVSWSP-DGEVIATGSMDNTIRLWD 215

QY     171 RCNKKGDFYRQNOISGAHN--TSDKOTP-----SKPKKKNSKGLAPSV----- 213
         :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db     216 KSGQCLG-----DAURG-HSKWI TSLSEWP IHLVKGPSKPRDASSKDGTTKIWDTVSRV 269

QY     214 -----DFQQSVTVVLFODENTLVLSAGADVGI IKVWDLR-----KNYTA YRQEPIA 258
         |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db     270 QOYTWMSGHTNSVCVKMGQGGLLYS-GSHDRTVRWVDINSQRCINILKSHAHVNNHLSL 328

QY     259 SKSFLY-----FCSTRKL--GYSSLILDSTGSTFLFANCTDDNIYFMNMTG 302
         |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db     329 STDVALRIGADHTGKPKSTPEEAOKKALENYEKICKNGNSEENMVTASDDYTMFLWN 388

QY     303 LK-TSPVAIPFNCHONSTFYVKSSLPDQFLVSGSSDEAAVIWK-----VSTPWQPPTV 355
         |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db     389 LKSTKPIARMYGHQKLVNHV--AFSPDGRIYVSASFDSNIKLDGRDGKFIST----- 439

QY     356 LLGHSQEVTSVCWCPSPDTKIATCSDDNLTAKIW--RLNRGLEEXPGDGKLSVGWASOKK 413
         |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db     440 FRGHIA SYVQAW-SSDCRELLVSCSKDTTLKWVDVTRFKLSVDLP EGIKTLYVDMSVDGK 498

QY      414 K 414
         :
Db      499 R 499

RESULT 9
US-09-291-170A-13
; Sequence 13, Application US/09291170A
; Patent No. 6410687
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; TITLE OF INVENTION: Depolymerization Inhibitors
; FILE REFERENCE: 18557B-000510US
; CURRENT APPLICATION NUMBER: US/09/291,170A
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 60/081,734
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Thermomonospora curvata
; FEATURE:
; OTHER INFORMATION: putative serine/threonine kinase PKWA WD40 repeat
; OTHER INFORMATION: region
US-09-291-170A-13

Query Match          6.2%; Score 241; DB 2; Length 251;
Best Local Similarity 23.7%; Pred. No. 2.2e-11;
Matches 85; Conservative 55; Mismatches 103; Indels 116; Gaps 14;

QY      35 SCNDEHTSYGET----GVFPVPFGCTFSSAPNMEHVLAVANEEGFVRLYNTESOSFRKKC 90
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       2 SGDELHTTEIGHTDMYRAVAFSPDGA-----LLASGSDDATVRLWDVAAAEER---- 48

```


193 ALLASGDDRTIRLMDVAAQ-EHHTTLEGHTEPVHSAFHP-EGTTLASASEGTIRIW 249

RESULT 11

```

RES001 11
US-09-724-592-13
; Sequence 13, Application US/09724592
; Patent No. 6699969
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald
; APPLICANT: Hartman, James
; TITLE OF INVENTION: Assays for the Detection of Microtubule Depolymerization Inhibitors
; FILE REFERENCE: UCSD-04884
; CURRENT APPLICATION NUMBER: US/09/724,592
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: PCT/US99/08086
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,734
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 251
; TYPE: prt
; ORGANISM: Thermomonospora curvata
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: putative serine/threonine kinase Pkwa WD40 repeat region
US-09-724-592-13

```

Query Match	6.2%	Score 241;	DB 2;	Length 251;
Best Local Similarity	23.7%	Pred. No. 2.2e-11;		
Matches	85;	Conservative	55;	Mismatches 103;
				Indels 116;
				Gaps 14;

Qy	35	SGNDEHTSYGET----	GVVPVPFGCTSSAPNMEHVLANVEBGFVRLYNTESQSFKKC	90
Db	2	SGDELHTLEIGHTDWRVAVFSDGA-----	LLASGSDDATVRLWDVAAAEER----	48
Qy	91	FKEWMAHNAVPLAWVPGEIKLVTAAAGDQATKFWDDVKAGELIGTCKGHQCSLKSAVFSK	15	
Db	49	-AVFEGHTHYVLVDIAFSPDGSVMVAGSGRSDGTARLNNVATGTEHAVLKGHTDYVVAVAFSP	10	
Qy	151	FEKAVFCTGRGDNIMWDTCRNKKDGFVRQVNOISGAHNSTDQTPSKPKKKQNSKGLA	21	
Db	108	DGSMVASGSRDGTIRLWDVATGKE-----	RDVLQ-----	13
Qy	211	PSVDFOQSPTVVLFDQENTLVSAGAVDGIKWDLRKNVTAYRQEPKIASKFLYPGSSTR	27	
Db	138	PA-----ENVVSLAFSPDGSMLVHGS--DSTVHLWD-----	VASGEALH-----	17
Qy	271	KLGYSLILDSTGSTLUFANCTDDNIYMFNMNTGLKTSPTVAIFNGHONSTFYVKS--SLSPDD	32	
Db	175	-----	TTEGH--TDWVRVAVAFSDG	19
Qy	330	QFLVSGSSDEAAVIWKVSTWMPQPTVVLHGHSQEVTSVCWCPSSDFTYKATCSDNTLKITW	388	
Db	193	ALLASGSDORTIRLWDVAQO-BEHTTLEGTPTSPHISVAFHP--EGTTLASASDGTIRIWI	249	

RESULT 12

RESULT 12
 US-09-673-222-13
 ; Sequence 13, Application US/09673222
 ; Patent No. 6872537
 ; GENERAL INFORMATION:
 ; APPLICANT: Vale, Ronald
 ; APPLICANT: Hartman, James
 ; TITLE OF INVENTION: Assays for the Detection of Microtubule Depolymerization Inhibitors
 ; FILE REFERENCE: UCSD-04765
 ; CURRENT APPLICATION NUMBER: US/09/673,222
 ; CURRENT FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: PCT/US99/08086
 ; PRIOR FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: 60/081,734

```

; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Thermomonospora curvata
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: putative serine/threonine kinase Pkwa WD40 repeat region
US-09-673-222-13

```

Query Match	6.2%;	Score 241;	DB 2;	Length 251;
Best Local Similarity	23.7%;	Pred. No. 2.2e-11;		
Matches	85;	Conservative 55;	Mismatches 103;	Indels 116; Gaps 14;
Qy	35	SGNDEHTSGET---	GVPPPPFGCTFSSAPNMEHVLAVANEGFVRLVNTESQSPFKKC	90
Db	2	SGDELHTEGHTDWWRAVAFSPDGA	-----LLASGDDATVRLMDVAAEER---	48
Qy	91	FEKEMAHNAVFDLAWVPGBELKLVTAAGDOTAKFPDVKAGELIGCTCKGHCOSLKSAFSK	150	
Db	49	-AVFEQHTHYVLDIASFPGSVMASGSRDGTARLNNVATCTEHAVLKGHTDYVVAFAFSP	107	
Qy	151	FEKAVFCTGRGDNIMWTRCNKKDGFYQVQNOISGAHNTSDKQTPSKPKKQNSKGLA	210	
Db	108	-DGSVMASGSRDGTIRLMDVATGKE---	RDVLQ-----A	137
Qy	211	PSVDFQGSVTVLFDQENTLVSAGVDGIKQVLDLRKNYTAYPEPIASKSFYPGSSSTR	270	
Db	138	PA-----ENVVSLAFSPDGSMLVHGS-DSTVHLMD-----	VASGEALH-----	174
Qy	271	KLGYSSILILDSTGTLFANCTDDNIYMFNMTGLKTSPPAIFNGHQNSTFYVKS-SLSPPDD	329	
Db	175	-----TFEGH-----	TDWRAVAFSPDG	192
Qy	330	QFLVSGSSDEAAIYKWTSPWPQPTVLLGHISQEVTSVCWCPSPFTTKIATCSDNDTLKIW	388	
Db	193	ALLASGSDRTIRLMDVAAO-EBHHTLEGTETPVSVAFHP-EQTTILASASEDGTIRIW	249	

RESULT 13

```

RESOL 13
US-08-883-534-3
; Sequence 3, Application US/0883534
; Patent No. 5846777
;
; GENERAL INFORMATION:
;
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
;
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
;
; NUMBER OF SEQUENCES: 6
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Diskette
;
; COMPUTER: IBM Compatible
;
; OPERATING SYSTEM: DOS
;
; SOFTWARE: FASSEQ for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/883,534
;
; FILING DATE:
;
; CLASSIFICATION: 424
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER:
;
; FILING DATE:
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Billings, Lucy J.
;
; REGISTRATION NUMBER: 36,749

```

```
; REFERENCE/DOCKET NUMBER: PF-0332 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: NEUTGWT01
; CLONE: 1221143
; US-08-883-534-3

Query Match          6.0%; Score 233.5; DB 1; Length 606;
Best Local Similarity 20.9%; Pred. No. 3.4e-10;
Matches 97; Conservative 79; Mismatches 157; Indels 131; Gaps 20;

QY 35 SGNDHTSYGETGVPVPPFGCTFSSAPNMEHV-----LAVANEFGVRLYNTES 83
Db 164 TGSDDNCAAFPEG---PPFKFPTVGDHSRFFVNCVRFPDGNRFATASADGQIYIY--DG 218
QY 84 QSFRRKC-PKEWMAHNAVFDLAWVPGELKLVTAAGDQTAKEFVDVKAGELIGT----- 135
Db 219 KTGEKVCALGGSKAHDGGIYAISWSPDTHLLSASGDKTSKIWDVSNVSVSTFPMGSTV 278
QY 136 -----CKGHQCSLKSVAFSK-PEKAVFC 157
Db 279 LDQQLGCLQKQDHLVSVLSGVINYLDNRNPKPLHVIKGHSKSIQCLTVHKGKSYIY 338
QY 158 TGRDGNIMVWTRCNKKGDFYRQ--VNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDF 215
Db 339 SGSHDGHINYWDSETGENDSFAGKGTNQVS--RMTVD-----ESGQLISCSMD 386
QY 216 QOSVTVLVFPD-----ENTLVSAG-----AVDGIHKVWDLRKNYT-----AY 252
Db 387 TVRYTSLMLRDYSGGVVVKLDVQPKCVAVPGGYAVVVCIGQIVLLKQDKQKCFSDINPGY 446
QY 253 ROEPIASKFLYPGSSTRKLG-----YSSL-----ILDSTGTLFANCTDDNI 295
Db 447 EPEVVA-----VHPGGDTVAIGVDGNVRLYSLITGLTKDEGKLEAKGPTVDVAYSHDGA 502
QY 296 YMFNMTGLKTSVP-AIFNGH-QNSTFY-----VKSSLSPPDQFLVSGSSDAAAYIKVYS 347
Db 503 FLAVCDASKVTVTFVSADGYSENNVPYGHAKIVCLAWSPDNEHFASGCGMDMMVYVWTLS 562
QY 348 TPWQPTVLLGHS-QEVTSCWCPCPSDFTKIATCSDDNLTAKIWRL 390
Db 563 DPETRVKIQDAHRLHHVSSLAWL--DEHTLVTTSHDASVKEWTI 604

RESULT 14
US-09-204-764-3
; Sequence 3, Application US/09204764
; Patent No. 6025464
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```


This Page Blank (uspto)

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:04:36 ; Search time 200 Seconds
(without alignments)
1668.840 Million cell updates/sec

Title: US-10-726-160-2

Perfect score: 3877

Sequence: 1 MUFNSVLRQPLQGLVLRNGWS.....TYFHRKSDQDFGCPHSTEL 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_8.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*
- 9: Geneseq2005s.*
- 10: Geneseq2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3877	100.0	730	6	ABU09611 Human ret
2	3877	100.0	730	7	Adf69740 Human ret
3	3877	100.0	730	8	ADO20069 Human PRO
4	3877	100.0	730	8	ADO20232 Human PRO
5	3877	100.0	730	9	Adx69661 Human ret
6	3877	100.0	730	9	Ady17438 PRO polytp
7	3877	100.0	730	10	Aee32574 Human L2D
8	3873	99.9	730	4	AAM25224 Human pro
9	3503.5	90.4	673	8	ADQ97697 Human can
10	3479.5	89.7	729	6	ABU09610 Mouse can
11	2913.5	75.1	786	8	ADQ97694 Mouse can
12	2293	59.1	434	4	AAB92779 Human pro
13	1439	37.1	332	4	ADM20092 Protein e
14	1413	36.4	287	4	ADM19846 Protein e
15	1216	31.4	232	4	AAB92715 Human pro
16	1165	30.0	522	8	ADO57307 Kidney de
17	1084	28.0	206	4	ABB50173 Human tra
18	1020	26.3	211	4	AB95316 Human pro
19	793.5	20.5	769	4	ABB65620 Drosophi
20	759	19.6	146	8	ADQ97699 Human can
21	587	15.1	523	8	ADT58375 Plant pol
22	406	10.5	111	4	AAG73994 Human col
23	312	8.0	610	8	ADS41755 Bacterial

24	298	7.7	1005	8	ADS41812	Ads41812 Bacterial
25	295.5	7.6	579	8	ADS29860	Ads29860 Bacterial
26	279.5	7.2	1118	8	ADS29812	Ads29812 Bacterial
27	269	6.9	608	8	ADS41794	Ads41794 Bacterial
28	267	6.9	1155	8	ADS30836	Ads30836 Bacterial
29	266	6.9	606	8	ADS29874	Ads29874 Bacterial
30	264.5	6.8	580	8	ADR86389	Adr86389 Aspergill
31	261.5	6.7	357	6	ABO53002	Abo53002 Human spl
32	261	6.7	478	8	ADS29872	Ads29872 Bacterial
33	259.5	6.7	357	4	AM93201	Aam93201 Human pol
34	259.5	6.7	357	8	ADL30556	Adl30556 Human pro
35	258.5	6.7	1117	8	ADS29932	Ads29932 Bacterial
36	256	6.6	348	9	ABE27318	Abe27318 Pinus rad
37	256	6.6	515	6	ABR53774	Abt53774 Protein s
38	256	6.6	515	7	ADK63846	Adk63846 Disease t
39	256	6.6	515	8	ADN18878	Adn18878 Bacterial
40	255.5	6.6	655	8	ADS41636	Ads41636 Bacterial
41	254.5	6.6	1136	8	ADS30013	Ads30013 Bacterial
42	254	6.6	379	8	ADN25500	Adn25500 Bacterial
43	253	6.5	375	4	AAE02854	Aae02854 Human GTP
44	253	6.5	375	5	AAE26083	Aae26083 Human GPA
45	253	6.5	375	6	ABG75809	Abg75809 Human GTP

ALIGNMENTS

RESULT 1

ABU09611
ID ABU09611 standard; protein; 730 AA.

XX

AC ABU09611;

XX

DT 01-JUL-2003 (first entry)

XX

DE Human retinoic acid-regulated nuclear matrix-associated protein.

XX

KW Nucleic acid identification; T cell receptor; T cell; tumour antigen;

KW

KW antigen presenting cell; HLA; nucleic acid library; ELISPOT assay;

KW

KW enzyme-linked immunospot; T cell receptor binding epitope;

KW

KW cancer associated antigen; ramp; human; polytope;

XX

OS retinoic acid-regulated nuclear matrix-associated protein.

XX

OS Homo sapiens.

XX

PN US2003003485-A1.

XX

PD 02-JAN-2003.

XX

PF 14-MAY-2002; 2002US-00145396.

XX

PR 15-MAY-2001; 2001US-0291125P.

 XX || PA | (LUDW-) LUDWIG INST CANCER RES. | |
PI	Uenaka A, Nakayama E;	
PI	WPI; 2003-416593/39.	
PI	N-PSDB; ABX95498.	
PT	Identifying nucleic acids encoding epitopes that bind T cell receptors on	
PT	T cells when presented by human leukocyte antigen by using an ELISPOT	
PT	assay that detects factors secreted by T cell in response to the receptor	
PT	binding.	
PS	Example 6; Fig 13; Sipp; English.	
PS	The invention describes a method of identifying nucleic acid encoding an	
PS	epitope that binds to a T cell receptor on T cell. The method involves	
PS	co-culturing antigen presenting cells that contain a nucleic acid library	
PS	and express HLA molecules presenting epitopes, with a T cell having a T	
PS	cell receptor that binds the epitope so that the T cell receptor binds	
PS	the epitope encoded by the nucleic acid library, detecting a factor	

CC secreted by the T cell in response to T cell receptor binding, using
CC ELISPOT (enzyme-linked immunospot) assay, and correlating factor
CC secretion with presence of a nucleic acid encoding the epitope. The
CC method is useful for identifying a nucleic acid molecule encoding the
CC epitope that specifically binds to a T cell receptor on a T cell when
CC presented by an HLA molecule. A second method described in the invention
CC is useful for identifying antigens that specifically bind to a T cell
CC receptor on a T cell. The method is useful for identifying cancer
CC associated antigens. This is the amino acid sequence of the human
CC retinoic acid-regulated nuclear matrix-associated protein (ramp) peptides
CC of which can be combined with peptides of one or more other cancer-
CC associated antigens to form polytopes
XX
SQ Sequence 730 AA;

```
Query Match      100.0%; Score 3877; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-293;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDHTSYGETGVPPFPFGCTFSSA 60
DB 1 MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDHTSYGETGVPPFPFGCTFSSA 60
QY 61 PMHEHLAVANEEGFVRLYNTESQSFRRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB 61 PMHEHLAVANEEGFVRLYNTESQSFRRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
QY 121 TAKFWDVKAGELIGTCKGHQCSLKSAFSEKFAVFCCTGGRDGNIMVMDTRCNKKGDFYR 180
DB 121 TAKFWDVKAGELIGTCKGHQCSLKSAFSEKFAVFCCTGGRDGNIMVMDTRCNKKGDFYR 180
QY 181 QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQOQSVTVVLFQDENTLVSAAGVDGII 240
DB 181 QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQOQSVTVVLFQDENTLVSAAGVDGII 240
QY 241 KWDLRKNYTAAYROEPIASKSFLYPGSSSTRKLGYSLLDSTGLFANCTDDNIMYFNM 300
DB 241 KWDLRKNYTAAYROEPIASKSFLYPGSSSTRKLGYSLLDSTGLFANCTDDNIMYFNM 300
QY 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIWKVSTPWPQPTVLLGHS 360
DB 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIWKVSTPWPQPTVLLGHS 360
QY 361 QEVTSVCMCPSPFTKIATCSDNTLKIWLRLNRLGEBKPGDGLKSTVGWASQKKESRPGL 420
DB 361 QEVTSVCMCPSPFTKIATCSDNTLKIWLRLNRLGEBKPGDGLKSTVGWASQKKESRPGL 420
QY 421 VVTSSQSTPAKAPRVKCNPSNPPSSSACAPSCAGDLPLPNTPTFSTKTSKARSP 480
DB 421 VVTSSQSTPAKAPRVKCNPSNPPSSSACAPSCAGDLPLPNTPTFSTKTSKARSP 480
QY 481 NRRGVSYSVPKPPSPFKMSIRNWVTRTPSSPPPTTPPASETAKIMSPPKALIPVSKSQSQ 540
DB 481 NRRGVSYSVPKPPSPFKMSIRNWVTRTPSSPPPTTPPASETAKIMSPPKALIPVSKSQSQ 540
QY 541 ABACSSRNVRKRRDLSSCLESVKQKCVKSCNVTLDQOVENHLDLCLAGNOEDLSK 600
DB 541 ABACSSRNVRKRRDLSSCLESVKQKCVKSCNVTLDQOVENHLDLCLAGNOEDLSK 600
QY 601 DSI GPTKSSKIECAGTSISEPPSPISPYASESGTLPPLRPGCGSEGMVGKENSPPENK 660
DB 601 DSI GPTKSSKIECAGTSISEPPSPISPYASESGTLPPLRPGCGSEGMVGKENSPPENK 660
QY 661 NWLLAAAKRKAENPSPSPSSQTPNSRRQSGKTLPSPTVITTPSSMRKICTYFHRKSQSD 720
DB 661 NWLLAAAKRKAENPSPSPSSQTPNSRRQSGKTLPSPTVITTPSSMRKICTYFHRKSQSD 720
QY 721 FCGPEHSTEL 730
DB 721 FCGPEHSTEL 730
```

ADF69740
ID ADF69740 standard; protein; 730 AA.
XX
AC ADF69740;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human retinoic acid (RA) regulated protein, 8.31.
XX
KW Human retinoic acid regulated gene; retinoic acid regulated protein; RA;
KW mitogen; morphogen; stem cell progenitor cell;
KW embryonic brain development; lung; liver; kidney;
KW cancer cell differentiation; Ushers Disease; cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003162268-A1.
XX
PD 28-AUG-2003.
XX
PF 08-APR-2003; 2003US-00409511.
XX
PR 14-JUL-1999; 99US-00354359.
XX
PA (IPNY/) IP N Y.
PA (CHEU/) W CHEUNG W M.
XX
PI Ip NY, Cheung WMW;
XX
DR WPI; 2003-897935/82.
DR N-PSDB; ADF69739.
XX
PT New retinoic acid regulated gene, useful in preparing a composition for
PT treating or preventing Ushers Disease or cancer.
XX
PS Claim 3; SEQ ID NO 2; 17pp; English.
XX
CC The present invention relates to the isolation of a novel human retinoic
CC acid (RA) regulated gene designated 8.31, and the polypeptide sequence it
CC encodes. The gene encoding the 8.31 protein maps to human chromosome
CC 1q32.1-32.2. The 8.31 protein is a mitogen or morphogen. It affects the
CC identity of stem cell progenitor cells or the development of embryonic
CC brain, lung, liver or kidney tissues. It also affects cancer cell
CC differentiation. The 8.31 protein is useful in treating or preventing
CC Ushers Disease or cancer. The present sequence represents human retinoic
CC acid (RA) regulated protein, 8.31.
XX
SQ Sequence 730 AA;

```
Query Match      100.0%; Score 3877; DB 7; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-293;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDHTSYGETGVPPFPFGCTFSSA 60
DB 1 MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDHTSYGETGVPPFPFGCTFSSA 60
QY 61 PMHEHLAVANEEGFVRLYNTESQSFRRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB 61 PMHEHLAVANEEGFVRLYNTESQSFRRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
QY 121 TAKFWDVKAGELIGTCKGHQCSLKSAFSEKFAVFCCTGGRDGNIMVMDTRCNKKGDFYR 180
DB 121 TAKFWDVKAGELIGTCKGHQCSLKSAFSEKFAVFCCTGGRDGNIMVMDTRCNKKGDFYR 180
QY 181 QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQOQSVTVVLFQDENTLVSAAGVDGII 240
DB 181 QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQOQSVTVVLFQDENTLVSAAGVDGII 240
QY 241 KWDLRKNYTAAYROEPIASKSFLYPGSSSTRKLGYSLLDSTGLFANCTDDNIMYFNM 300
DB 241 KWDLRKNYTAAYROEPIASKSFLYPGSSSTRKLGYSLLDSTGLFANCTDDNIMYFNM 300
```


QY 301 TGLKTSPTVAIFNGHQNSTFYVKSSLSPPDQFLVSGSSDEAAIWKVSTPWQPTVLLGHS 360
DB 301 TGLKTSPTVAIFNGHQNSTFYVKSSLSPPDQFLVSGSSDEAAIWKVSTPWQPTVLLGHS 360
QY 361 QEVTSVCWCPSDFTKIATCSDNTLKIWLNRGLKEKPGDKLSTVGWASQKKGSRPGL 420
DB 361 QEVTSVCWCPSDFTKIATCSDNTLKIWLNRGLKEKPGDKLSTVGWASQKKGSRPGL 420
QY 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480
DB 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480
QY 481 NRRGSVSSVPKPPSSFKMSIRNWVTRTPSSPPITPPASETKIMSPRKALIPVSKSSQ 540
DB 481 NRRGSVSSVPKPPSSFKMSIRNWVTRTPSSPPITPPASETKIMSPRKALIPVSKSSQ 540
QY 541 AEACSESRNRVKRRLDSSCLSVKQCVKSCNCVTELDQOVENHLDLCLAGNQEDLSK 600
DB 541 AEACSESRNRVKRRLDSSCLSVKQCVKSCNCVTELDQOVENHLDLCLAGNQEDLSK 600
QY 601 DSLGPTKSKIEGAGTSISEPPSPISPYASESGTLPLPLPCGEGSEWVGKENSSENK 660
DB 601 DSLGPTKSKIEGAGTSISEPPSPISPYASESGTLPLPLPCGEGSEWVGKENSSENK 660
QY 661 NWLLAMAARKAENPSPRSPSSQTNRRSQSGKTLPSPTIIPSSMRKICTYFHRKSQED 720
DB 661 NWLLAMAARKAENPSPRSPSSQTNRRSQSGKTLPSPTIIPSSMRKICTYFHRKSQED 720
QY 721 FCGPEHSTEL 730
DB 721 FCGPEHSTEL 730
RESULT 3
ADO20069
ID ADO20069 standard; protein; 730 AA.
AC ADO20069;
XX
DT 12-AUG-2004 (first entry)
DE Human PRO polypeptide #489.
KW Human; PRO: immune related disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
KW diabetes mellitus; renal disease; demyelinating disease;
KW central nervous system; peripheral nervous system;
KW demyelinating polyneuropathy; Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX
OS Homo sapiens.
XX WO2004043361-A2.
PN
XX 27-MAY-2004.
PD
XX
PF 06-NOV-2003; 2003WO-US035268..
XX
PR 08-NOV-2002; 2002US-0425235P.
XX
XX (GETH) GENENTECH INC.
XX
PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PW;
PI Wood WJ, Wu TD;
XX
XX WPI; 2004-420067/39.
DR N-PSDB; ADO20068.
XX
PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or

PT spondyloarthropathy.
PS Claim 7; SEQ ID NO 978; 1731pp; English.
XX
CC The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polypeptide of the
CC invention.
XX
SQ Sequence 730 AA;
Query Match 100.0%; Score 3877; DB 8; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-293;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLFNSVLRLPOLGVLNRGWSQYPLQSLLTGYQCSGNDHTSYGETGVPPVPPGCTFSSA 60
DB 1 MLFNSVLRLPOLGVLNRGWSQYPLQSLLTGYQCSGNDHTSYGETGVPPVPPGCTFSSA 60
QY 61 PMEHVLAVANEHGFVRLYNTESQSFRRKCKFKNWAHNAVFDLAWVPELKLVTAAQGO 120
DB 61 PMEHVLAVANEHGFVRLYNTESQSFRRKCKFKNWAHNAVFDLAWVPELKLVTAAQGO 120
QY 121 TAKPMDVRAGELIGTCKGHQCSLSKSAFVKFAVCTGGRDGNIMWDTNCKKDGFPYR 180
DB 121 TAKPMDVRAGELIGTCKGHQCSLSKSAFVKFAVCTGGRDGNIMWDTNCKKDGFPYR 180
QY 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDPQQSVTVVLFODENTLVSAGAVDGI 240
DB 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDPQQSVTVVLFODENTLVSAGAVDGI 240
QY 241 KWDLRKNTAYRQEPPIASKSFLYPGSSTRKLGYSLLDSTGLFANCTDDNLYMFNM 300
DB 241 KWDLRKNTAYRQEPPIASKSFLYPGSSTRKLGYSLLDSTGLFANCTDDNLYMFNM 300
QY 301 TGLKTSPTVAIFNGHQNSTFYVKSSLSPPDQFLVSGSSDEAAIWKVSTPWQPTVLLGHS 360
DB 301 TGLKTSPTVAIFNGHQNSTFYVKSSLSPPDQFLVSGSSDEAAIWKVSTPWQPTVLLGHS 360
QY 361 QEVTSVCWCPSDFTKIATCSDNTLKIWLNRGLKEKPGDKLSTVGWASQKKGSRPGL 420
DB 361 QEVTSVCWCPSDFTKIATCSDNTLKIWLNRGLKEKPGDKLSTVGWASQKKGSRPGL 420
QY 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480
DB 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480
QY 481 NRRGSVSSVPKPPSSFKMSIRNWVTRTPSSPPITPPASETKIMSPRKALIPVSKSSQ 540
DB 481 NRRGSVSSVPKPPSSFKMSIRNWVTRTPSSPPITPPASETKIMSPRKALIPVSKSSQ 540
QY 541 AEACSESRNRVKRRLDSSCLSVKQCVKSCNCVTELDQOVENHLDLCLAGNQEDLSK 600
DB 541 AEACSESRNRVKRRLDSSCLSVKQCVKSCNCVTELDQOVENHLDLCLAGNQEDLSK 600
QY 601 DSLGPTKSKIEGAGTSISEPPSPISPYASESGTLPLPLPCGEGSEWVGKENSSENK 660
DB 601 DSLGPTKSKIEGAGTSISEPPSPISPYASESGTLPLPLPCGEGSEWVGKENSSENK 660
QY 661 NWLLAMAARKAENPSPRSPSSQTNRRSQSGKTLPSPTIIPSSMRKICTYFHRKSQED 720
DB 661 NWLLAMAARKAENPSPRSPSSQTNRRSQSGKTLPSPTIIPSSMRKICTYFHRKSQED 720
QY 721 FCGPEHSTEL 730
DB 721 FCGPEHSTEL 730

Db 721 FCGPEHSTEL 730

RESULT 4

ID ADO20232 standard; protein; 730 AA.

XX ADO20232;

XX 12-AUG-2004 (first entry)

DT Human PRO polypeptide #567.

DE

XX Human; PRO; immune related disorder; systemic lupus erythematosus;

XX rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;

KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;

KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;

KW diabetes mellitus; renal disease; demyelinating disease;

KW central nervous system; peripheral nervous system;

KW demyelinating polyneuropathy; Guillain-Barre syndrome;

KW chronic inflammatory demyelinating polyneuropathy.

XX Homo sapiens.

OS

XX WO2004043361-A2.

PN

XX 27-MAY-2004.

PD

XX 06-NOV-2003; 2003WO-US035268.

PF

XX 08-NOV-2002; 2002US-0425235P.

PR

XX (GETH) GENENTECH INC.

PA

PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PW;

PI Wood WI, Wu TD;

XX WPI; 2004-420067/39.

DR N-PSDB; ADO20231.

DR

XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for

PT treating an immune related disorder such as systemic lupus erythematosus,

PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or

PT spondyloarthropathy.

PT

XX Claim 7; SEQ ID NO 1134; 1731pp; English.

PS

XX The invention relates to human PRO polypeptides and the polynucleotides

CC encoding them. The polypeptides and polynucleotides are useful for

CC treating and diagnosing immune related disorders in mammals. The immune

CC related disorders include systemic lupus erythematosus, rheumatoid

CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic

CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune

CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes

CC mellitus, immune-mediated renal disease, demyelinating diseases of the

CC central or peripheral nervous system, demyelinating polyneuropathy,

CC Guillain-Barre syndrome and chronic inflammatory demyelinating

CC polyneuropathy. This sequence represents a human PRO polypeptide of the

CC invention.

XX

XX Sequence 730 AA;

Query Match 100.0%; Score 3877; DB 8; Length 730;

Best Local Similarity 100.0%; Pred. No. 1.4e-293;

Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFNSVLROQLGVLNGLSSQYPLQSLTGTQCSGNDEHTSYGETGVPPFPFGCTFSSA 60

DB 1 MLFNSVLROQLGVLNGLSSQYPLQSLTGTQCSGNDEHTSYGETGVPPFPFGCTFSSA 60

QY 61 PNMEHVLAVANEGFVRLNVTESQSRKCFKEWMAHNAVFDLAWVPGELKLVTAAGDQ 120

DB 61 PNMEHVLAVANEGFVRLNVTESQSRKCFKEWMAHNAVFDLAWVPGELKLVTAAGDQ 120

QY 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTFCNKKDGFYR 180

DB 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTFCNKKDGFYR 180

QY 181 QVNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGI 240

DB 181 QVNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGI 240

QY 241 KVDLRLKNYTAIROEPIASKSFLYPGSSTRKLGYSLLDSTGLFANCTDDNIYFNM 300

DB 241 KVDLRLKNYTAIROEPIASKSFLYPGSSTRKLGYSLLDSTGLFANCTDDNIYFNM 300

QY 301 TGLKTSVAIFNGHQNSTFYVKSSISPDQPLVSGSSDEAAIYWKVSTPWOPPTVLLGHS 360

DB 301 TGLKTSVAIFNGHQNSTFYVKSSISPDQPLVSGSSDEAAIYWKVSTPWOPPTVLLGHS 360

QY 361 QEVTSVCWCPSDFTKIATCSDNTLKIWRLNRGLEEKPGDKLSTVGWASOKKESRPG 420

DB 361 QEVTSVCWCPSDFTKIATCSDNTLKIWRLNRGLEEKPGDKLSTVGWASOKKESRPG 420

QY 421 VVTSSQSTPAKAPRVKCNPSNPSSSAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480

DB 421 VVTSSQSTPAKAPRVKCNPSNPSSSAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480

QY 481 NRGSVSVSPKPPSSFKMSIRNWTRTPSSSPPTTPPASETKIMSPRKALIPVSQSSQ 540

DB 481 NRGSVSVSPKPPSSFKMSIRNWTRTPSSSPPTTPPASETKIMSPRKALIPVSQSSQ 540

QY 541 AEACSESRNRKRLDSSCLESVKQKVCVSCNCTELDQVENLHLDLCCLAGNQEDLSK 600

DB 541 AEACSESRNRKRLDSSCLESVKQKVCVSCNCTELDQVENLHLDLCCLAGNQEDLSK 600

QY 601 DSLGPTKSSKIEGAGTISIEPPSPISPYASESCGTPLPLRPGCGESBMVGKENSSENK 660

DB 601 DSLGPTKSSKIEGAGTISIEPPSPISPYASESCGTPLPLRPGCGESBMVGKENSSENK 660

QY 661 NWLLAMAAKRAENPSPRSPSSQTPNSRRQSGKTLPSVTTTPSSMRKICTYFHRKSD 720

DB 661 NWLLAMAAKRAENPSPRSPSSQTPNSRRQSGKTLPSVTTTPSSMRKICTYFHRKSD 720

QY 721 FCGPEHSTEL 730

DB 721 FCGPEHSTEL 730

RESULT 5

ADK69661

ID ADK69661 standard; protein; 730 AA.

XX

AC ADK69661;

XX

DT 05-MAY-2005 (first entry)

XX

DE Human retinoic acid regulated nuclear matrix protein.

XX

KW DNA purification; retinoic acid; diagnosis; hepatocellular carcinoma;

KW cytostatic; prognosis; cancer; liver tumor;

KW retinoic acid regulated nuclear matrix protein.

XX

OS Homo sapiens.

XX

PN US2005037372-A1.

XX

PD 17-FEB-2005.

XX

PF 02-DEC-2003; 2003US-00726160.

XX

PR 14-JUL-1999; 99US-00354359.

PR 08-APR-2003; 2003US-00409511.

XX

PA (UYHK-) UNIV HONG KONG SCI & TECHNOLOGY.

XX

```
PI Ip NY, Cheung WW;
XX WPI: 2005-172246/18.
DR N-PSDB; ADX69660.
XX Novel isolated retinoic acid regulated gene, useful for screening and
PT determining prognosis of patient having Hepatocellular cancer.
XX Claim 2; SEQ ID NO 2; 21pp; English.
XX
XX The invention relates to an isolated retinoic acid regulated gene (I)
CC having a fully defined 281 nucleotides sequence given in specification.
CC A protein (II) encoded by (I) is useful as a screening tool for
CC diagnosing Hepatocellular carcinomas and for monitoring treatment or
CC progression of Hepatocellular carcinomas. An antibody (III) to (II) is
CC useful for screening and determining the prognosis of a patient having
CC Hepatocellular cancer, where the biological samples comprise liver
CC tissues. This sequence corresponds to the retinoic acid regulated nuclear
CC matrix protein (RAMP; II).
XX
SQ Sequence 730 AA;

Query Match 100.0%; Score 3877; DB 9; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-293;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFNSVLROPQLGVLNRGWSQYPLQSLITGYOCSGNDEHTSYGETGVPPPGCTFSSA 60
DB 1 MLFNSVLROPQLGVLNRGWSQYPLQSLITGYOCSGNDEHTSYGETGVPPPGCTFSSA 60
QY 61 PNMEHLVANEEGFVRLYNTESQSFRRKCKFKEWMAHNAVFDLAWVPGLKLVTAAGDQ 120
DB 61 PNMEHLVANEEGFVRLYNTESQSFRRKCKFKEWMAHNAVFDLAWVPGLKLVTAAGDQ 120
QY 121 TAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTRCNKKDGYR 180
DB 121 TAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTRCNKKDGYR 180
QY 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSQSVTVLFDQENTLVSAGVDGII 240
DB 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSQSVTVLFDQENTLVSAGVDGII 240
QY 241 KWDLRKNYTAIROEPIASKSFLYPGSSTRKGLYSSLLDSTGTLFANCTDDNIMVFNM 300
DB 241 KWDLRKNYTAIROEPIASKSFLYPGSSTRKGLYSSLLDSTGTLFANCTDDNIMVFNM 300
QY 301 TGLKTSVAIFNGHONSTFYVYKSSLSDDQFLVSGSSDEAAYIWKVSTPQPTVLLGHS 360
DB 301 TGLKTSVAIFNGHONSTFYVYKSSLSDDQFLVSGSSDEAAYIWKVSTPQPTVLLGHS 360
QY 361 QEVTSVCWCPDFTFKIATCSDDNLTIKIWLNRGLERKPGDKLSTVGSWASQKKESRPGI 420
DB 361 QEVTSVCWCPDFTFKIATCSDDNLTIKIWLNRGLERKPGDKLSTVGSWASQKKESRPGI 420
QY 421 VVTSSOSTPAKAPRVKCNPSNPSAAACAPSAGDLPLPSNTPPTSIIKTSAPAKARSP 480
DB 421 VVTSSOSTPAKAPRVKCNPSNPSAAACAPSAGDLPLPSNTPPTSIIKTSAPAKARSP 480
QY 481 NRRGSVSVSPKPPSSFKMIRNMTVTRTPSSPPIPTPASETKIMSPRKALIPVSQKSSQ 540
DB 481 NRRGSVSVSPKPPSSFKMIRNMTVTRTPSSPPIPTPASETKIMSPRKALIPVSQKSSQ 540
QY 541 AEACSESRNRVKRLDSSCLSVKQKVCNCVTELDGQVENLHLDLCLAGNQEDLSK 600
DB 541 AEACSESRNRVKRLDSSCLSVKQKVCNCVTELDGQVENLHLDLCLAGNQEDLSK 600
QY 601 DSLGPTKSSKLEGAGTSGISRPSPISPVASESCGTLPLRPPCGEGSEMGKENSPEK 660
DB 601 DSLGPTKSSKLEGAGTSGISRPSPISPVASESCGTLPLRPPCGEGSEMGKENSPEK 660
QY 661 NWLLAMAARKAENPSPRSPSSQTPNRRSGKTLPSFVTITPSSMRKICTYFHRKSQED 720
DB 661 NWLLAMAARKAENPSPRSPSSQTPNRRSGKTLPSFVTITPSSMRKICTYFHRKSQED 720

QY 721 FCGPEHSTEL 730
DB 721 FCGPEHSTEL 730

RESULT 6
ADY17438
ID ADY17438 standard; protein; 730 AA.
XX AC ADY17438;
XX 05-MAY-2005 (first entry)
XX PRO polypeptide SEQ ID NO 3244.
XX Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotrophic;
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW Antiallergic; ds; gene; diagnosis.
XX
XX Homo sapiens.
OS
XX WO2005016962-A2.
XX
XX 24-FEB-2005.
XX
XX 11-AUG-2004; 2004WO-US026249.
XX
XX 11-AUG-2003; 2003US-0493546P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
XX WPI; 2005-182330/19.
XX
XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
XX Claim 8; SEQ ID NO 3244; 158pp; English.
XX
XX The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC composition, and method are useful for diagnosing and treating an immune
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC arthritis. The present sequence represents a DNA encoding a PRO
CC polypeptide.
XX
XX Sequence 730 AA;

Query Match 100.0%; Score 3877; DB 9; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-293;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFNSVLROPQLGVLNRGWSQYPLQSLITGYOCSGNDEHTSYGETGVPPPGCTFSSA 60
DB 1 MLFNSVLROPQLGVLNRGWSQYPLQSLITGYOCSGNDEHTSYGETGVPPPGCTFSSA 60
QY 61 PNMEHLVANEEGFVRLYNTESQSFRRKCKFKEWMAHNAVFDLAWVPGLKLVTAAGDQ 120
DB 61 PNMEHLVANEEGFVRLYNTESQSFRRKCKFKEWMAHNAVFDLAWVPGLKLVTAAGDQ 120
QY 121 TAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTRCNKKDGYR 180
DB 121 TAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTRCNKKDGYR 180
QY 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSQSVTVLFDQENTLVSAGVDGII 240
DB 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSQSVTVLFDQENTLVSAGVDGII 240
```

QY 241 KWDLRKNYAYROEPIASKSFLYPGSTRKGLGYSLLDSTGLFANCTDDNLYMFNM 300
DB |||||||
QY 241 KWDLRKNYAYROEPIASKSFLYPGSTRKGLGYSLLDSTGLFANCTDDNLYMFNM 300
DB |||||||
QY 301 TGLKTSPIVAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB |||||||
QY 301 TGLKTSPIVAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB |||||||
QY 361 QEVTSVCWCPDFTKIATCSDNTLKIWLRNGLBEKPGDKLSTVGWASQKKESRPGI 420
DB |||||||
QY 361 QEVTSVCWCPDFTKIATCSDNTLKIWLRNGLBEKPGDKLSTVGWASQKKESRPGI 420
DB |||||||
QY 421 VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPSNTFTFSIKTSPAKARSP 480
DB |||||||
QY 421 VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPSNTFTFSIKTSPAKARSP 480
DB |||||||
QY 481 NRRGVSSVSPKPPSPFKMSIRNWTRTPSSPPPTPPASETKIMSPKALIPVQKSSQ 540
DB |||||||
QY 481 NRRGVSSVSPKPPSPFKMSIRNWTRTPSSPPPTPPASETKIMSPKALIPVQKSSQ 540
DB |||||||
QY 541 ABACSESRRVRRLLDSSCLESVKQKVKSCNCVTLDQOVENLHDLCLLAGNQEDLSK 600
DB |||||||
QY 601 DSLGPTKSSKIGAGTSSISEPSPISPYASESCGTLPLLRPCRGSEMVKGKSSPENK 660
DB |||||||
QY 601 DSLGPTKSSKIGAGTSSISEPSPISPYASESCGTLPLLRPCRGSEMVKGKSSPENK 660
DB |||||||
QY 661 NWLLAAAKRKAENSPRSPSSQTPNSRQSGKTLPSPTIIPSSMRKICTYFHRKSQED 720
DB |||||||
QY 721 FCGPEHSTEL 730
DB |||||||
QY 721 FCGPEHSTEL 730
DB |||||||

RESULT 7
ID AEE32574
XX AEE32574 standard; protein; 730 AA.

AC AEE32574;
DT 09-FEB-2006 (first entry)
DE Human L2DTL protein SEQ ID NO 8.
KW cell cycle; p53; cancer; neoplasm; cytotstatic; lung tumor; breast tumor;
KW ovarian tumor; ubiquitin; L2DTL.

OS Homo sapiens.
XX WO2005114188-A2.
PN 01-DEC-2005.
XX 27-APR-2005; 2005WO-US014615.
XX 27-APR-2004; 2004US-0565707P.

XX (UYVA) UNIV YALE.
XX Zhang H, Banks DP, Higa LAA;
XX WPI: 2006-020151/02.
XX GENBANK; AAF35182.
XX Identifying compound for modulating ubiquitination of p53 polypeptide by
PT incubating CUL4 E3 ligase complex with a p53 polypeptide and comparing
PT ubiquitination of p53 polypeptide in the presence or absence of compound.

XX Disclosure; SEQ ID NO 8; 88pp; English.

CC This invention describes a novel method for identifying a compound which
CC modulates the ubiquitination of a p53 polypeptide. This method involves
CC incubating the CUL4 E3 ligase complex (comprising CUL4A, CUL4B, DBB1,
CC L2DTL, ROC1, ROC2, MDM2 and Pih2, preferably Pih2) with a p53
CC polypeptide for the ubiquitination of p53 and comparing the
CC ubiquitination of the p53 polypeptide measured in the step of incubating
CC the test compound to determine a difference in the ubiquitination of the
CC p53 substrate. The ligase complex is incubated with a p53 polypeptide in
CC the presence of one or more of ubiquitin-activating E1 enzymes, ubiquitin
CC -conjugating E2 enzymes, ATPs and ubiquitins. The ubiquitin-conjugating
CC E2 enzyme is chosen from Cdc34, UbcH1, UbcH2, UbcH3, UbcH4, UbcH5, UbcH6,
CC UbcH7, UbcH10 and L-UBC. The UbcH5 is UbcH5A, UbcH5B or UbcH5C. Methods
CC are also disclosed for a) decreasing the level of a p53 polypeptide in a
CC cell, by increasing the expression level of a CUL4-associated
CC polynucleotide (e.g. an E2 conjugation enzymes chosen from CDC34 E2,
CC UbcH5B E2, CUL4A, CUL4B, DBB1, L2DTL, ROC1, ROC2, MDM2, Pih2 and UbcH5C
CC E2) in the cell, increasing the level of a CUL4-associated polypeptide in
CC the cell, increasing binding between a p53 polypeptide and a CUL4 E3
CC ligase complex, increasing the ubiquitination activity of a CUL4 ligase
CC complex, increasing the binding between a CUL4 ligase complex and MDM2,
CC increasing the binding between a CUL4 ligase complex and Pih2, or
CC increasing the activity of CUL1 E3 ligase complex and b) increasing the
CC level of a p53 polypeptide in a cell, which involves decreasing the
CC expression level of an mRNA encoding a CUL4-associated polypeptide in the
CC cell, decreasing the level of a CUL4-associated polypeptide in the cell,
CC decreasing binding between a p53 polypeptide and a CUL4 E3 ligase
CC complex, decreasing the ubiquitination activity of a CUL4 ligase complex,
CC decreasing the binding between a CUL4 ligase complex and MDM2, decreasing
CC the binding between a CUL4 ligase complex and Pih2 or decreasing the
CC activity of a CUL1 E3 ligase complex. The methods of the invention are
CC useful for identifying modulators of ubiquitination of p53 polypeptide
CC utilized for treating cancer e.g. lung cancer, breast cancer or ovarian
CC carcinoma and enabling a high-throughput identification method for
CC modulators of ubiquitination of p53 polypeptide. This sequence represents
CC the human L2DTL protein, part of the CUL4 E3 ligase complex incubated
CC with p53 to investigate ubiquitination.

XX SQ Sequence 730 AA;

Query Match 100.0%; Score 3877; DB 10; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-293;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFNSVLRLQPOLGVLNRNGWSSQYPLQSLLTGYQCSGNDHSTSYGETGVPPVPGCTFSSA 60
DB |||||||
QY 1 MLFNSVLRLQPOLGVLNRNGWSSQYPLQSLLTGYQCSGNDHSTSYGETGVPPVPGCTFSSA 60
DB |||||||
QY 61 PMEHVLAIVANEPEGFVRLYNTESQSFRRKKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB |||||||
QY 61 PMEHVLAIVANEPEGFVRLYNTESQSFRRKKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB |||||||
QY 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKPEKAVFCTGGRDGNIMVMDTRCNCKDGFYR 180
DB |||||||
QY 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKPEKAVFCTGGRDGNIMVMDTRCNCKDGFYR 180
DB |||||||
QY 181 QVNIQISGAHNTSDKQTPSKPKKQNSKGLAPSDVDFQQSVTVVVLFDENTLVSAGAVDGI 240
DB |||||||
QY 181 QVNIQISGAHNTSDKQTPSKPKKQNSKGLAPSDVDFQQSVTVVVLFDENTLVSAGAVDGI 240
DB |||||||
QY 241 KWDLRKNYAYROEPIASKSFLYPGSTRKGLGYSLLDSTGLFANCTDDNLYMFNM 300
DB |||||||
QY 241 KWDLRKNYAYROEPIASKSFLYPGSTRKGLGYSLLDSTGLFANCTDDNLYMFNM 300
DB |||||||
QY 301 TGLKTSPIVAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB |||||||
QY 301 TGLKTSPIVAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB |||||||
QY 361 QEVTSVCWCPDFTKIATCSDNTLKIWLRNGLBEKPGDKLSTVGWASQKKESRPGI 420
DB |||||||
QY 361 QEVTSVCWCPDFTKIATCSDNTLKIWLRNGLBEKPGDKLSTVGWASQKKESRPGI 420
DB |||||||
QY 421 VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPSNTFTFSIKTSPAKARSP 480
DB |||||||

Db 421 VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480
Qy 481 NRRGSVSVSPKPPSFKSIRNWTTRTPSSSPPIPPASETKIMSPKALIPVSQKSSQ 540
Db 481 NRRGSVSVSPKPPSFKSIRNWTTRTPSSSPPIPPASETKIMSPKALIPVSQKSSQ 540
Qy 541 AEACSESRNRVKRRLLDSSCLSVKQKVCNVCNVTDLGQVENLHLDLCLAGNQBDLSK 600
Db 541 AEACSESRNRVKRRLLDSSCLSVKQKVCNVCNVTDLGQVENLHLDLCLAGNQBDLSK 600
Qy 601 DSLGPTKSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPGEGSEVMVGKENSSENK 660
Db 601 DSLGPTKSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPGEGSEVMVGKENSSENK 660
Qy 661 NWLLAAAKRAENPSPRSPSSQTPNRRQSGTKLPSPTTTPSSMRKICTYFHRKSQED 720
Db 661 NWLLAAAKRAENPSPRSPSSQTPNRRQSGTKLPSPTTTPSSMRKICTYFHRKSQED 720
Qy 721 FCGPEHSTEL 730
Db 721 FCGPEHSTEL 730

RESULT 8

AA25224
ID AAM25224 standard; protein; 730 AA.
XX AAM25224;
AC AAM25224;
XX AAM25224;
DT 04-DEC-2001 (first entry)
XX Human prostate-related gene 83P5G4 encoded protein.

83P5G4-related protein; prostate; testis; tissue; cancer; bladder; liver;
tumour; kidney; brain; bone; ovary; breast; pancreas; colon; lung; serum;
cytostatic; gene therapy; antibody therapy; ribozyme; blood; cervix;
single chain monoclonal antibody; urine; uterus; rectum; stomach; human;
chromosome 1q31-q32.

XX Homo sapiens.

XX WO200159115-A2.

XX 16-AUG-2001.

XX 09-FEB-2001; 2001WO-US004426.

XX 09-FEB-2000; 2000US-0181261P.

XX (UROC-) UROGENESYS INC.

XX Hubert RS, Afar DEH, Challita-Eid PM, Paris M, Levin E;

PI Mitchell SC, Jakobovits A;

XX WPI; 2001-514669/56.

DR N-PSDB; AAR99153.

PT An isolated 83P5G4-related protein useful as a diagnostic and/or
therapeutic agent in multiple cancers such as prostate, bladder and bone
cancer.

XX Claim 12; Fig 2; 112pp; English.

XX The polypeptide sequences represent the 83P5G4-related protein and
CC peptide fragments of the protein. 83P5G4 exhibits prostate specific
CC expression in normal adult tissue, but it is also aberrantly expressed in
CC many cancers including tumours of the prostate, testis, bladder, kidney,
CC brain, bone, cervix, uterus, ovary, breast, pancreas, stomach, rectum,
CC liver, colon and lung. The 83P5G4 polynucleotide, its related protein and
CC peptide fragments and specific PCR primers are therefore useful for
CC diagnosing and treating cancer. A vector comprising a polynucleotide
CC which encodes a single chain monoclonal antibody, that immunospecifically

CC binds to an 83P5G4-related protein, and a ribozyme capable of cleaving a
CC polynucleotide having the 83P5G4 coding sequence, are both useful in the
CC preparation of a composition for treating a patient with a cancer that
CC expresses 83P5G4. The sequences can be used in diagnostic methods to
CC monitor the level of 83P5G4 gene products in serum, blood, urine and
CC tissue and to thereby detect the presence of cancerous cells
XX

SQ Sequence 730 AA;

Query Match 99.9%; Score 3873; DB 4; Length 730;
Best Local Similarity 99.9%; Pred. No. 2.9e-293;
Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLFNSVLROPOLGVLNRGWSQYPLQSLLTGYQCGNDEHTSYGETGVVPVPGCTFSSA 60

Db 1 MLFNSALRQPOLGVLNRGWSQYPLQSLLTGYQCGNDEHTSYGETGVVPVPGCTFSSA 60

Qy 61 PMEHVLAVANEEGVRLYNTESQSFRRKCFKEMAHNAVFDLAWVPGLKLVTAAGDQ 120

Db 61 PMEHVLAVANEEGVRLYNTESQSFRRKCFKEMAHNAVFDLAWVPGLKLVTAAGDQ 120

Qy 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKPKAVFCTGGRDGNIMVWTRCNKKGDFYR 180

Db 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKPKAVFCTGGRDGNIMVWTRCNKKGDFYR 180

Qy 181 QVNOISGAHNTSDKOTPSKPKKQNSKGLAPSDVDFQOSVTVVLFOEDNTLVSAGAVDGI 240

Db 181 QVNOISGAHNTSDKOTPSKPKKQNSKGLAPSDVDFQOSVTVVLFOEDNTLVSAGAVDGI 240

Qy 241 KVDLRKNYTAAYROBPIASKSFYPCSSSTRKLGYSLLDSTGSTLFANCTDDNIYMFNM 300

Db 241 KVDLRKNYTAAYROBPIASKSFYPCSSSTRKLGYSLLDSTGSTLFANCTDDNIYMFNM 300

Qy 301 TGLKTSPIVAIFNGHONSTFYVKSLSLSPDDQFLVSSGSDDEAAIYKWTSTPQPTVLLGHS 360

Db 301 TGLKTSPIVAIFNGHONSTFYVKSLSLSPDDQFLVSSGSDDEAAIYKWTSTPQPTVLLGHS 360

Qy 361 QSVTSVCWCPSDFTKIATCDDNTLKIWLNRGLEEKPGDKLSTVGWASQKKESRPG 420

Db 361 QSVTSVCWCPSDFTKIATCDDNTLKIWLNRGLEEKPGDKLSTVGWASQKKESRPG 420

Qy 421 VTVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480

Db 421 VTVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480

Qy 481 NRRGSVSVSPKPPSPFKMSIRNWTTRTPSSSPPIPPASETKIMSPKALIPVSQKSSQ 540

Db 481 NRRGSVSVSPKPPSPFKMSIRNWTTRTPSSSPPIPPASETKIMSPKALIPVSQKSSQ 540

Qy 541 AEACSESRNRVKRRLLDSSCLSVKQKVCNVCNVTDLGQVENLHLDLCLAGNQBDLSK 600

Db 541 AEACSESRNRVKRRLLDSSCLSVKQKVCNVCNVTDLGQVENLHLDLCLAGNQBDLSK 600

Qy 601 DSLGPTKSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPGEGSEVMVGKENSSENK 660

Db 601 DSLGPTKSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPGEGSEVMVGKENSSENK 660

Qy 661 NWLLAAAKRAENPSPRSPSSQTPNRRQSGTKLPSPTTTPSSMRKICTYFHRKSQED 720

Db 661 NWLLAAAKRAENPSPRSPSSQTPNRRQSGTKLPSPTTTPSSMRKICTYFHRKSQED 720

Qy 721 FCGPEHSTEL 730

Db 721 FCGPEHSTEL 730

RESULT 9

ADQ97697

ID ADQ97697 standard; protein; 673 AA.

XX ADQ97697;

XX 07-OCT-2004 (first entry)

XX Human cancer associated sequence HP1-10-029, SEQ ID 674.
DE Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human.
KW Homo sapiens.
OS WO2004060304-A2.
PN 22-JUL-2004.
PD 22-DEC-2003; 2003WO-US041389.
PF 27-DEC-2002; 2002US-00330773.
PR (SAGR-) SAGRES DISCOVERY INC.
XX Morris DW, Malandro MS;
PI WPI; 2004-543781/52.
DR New isolated cancer associated nucleic acids comprising at least 10
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
PT Claim 1; SEQ ID NO 674; 199pp; English.
XX The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 673 AA;
SQ
Query Match 90.4%; Score 3503.5; DB 8; Length 673;
Best Local Similarity 91.8%; Pred. No. 26-264; 2; Indels 57; Gaps 2;
Matches 670; Conservative 1; Mismatches 2;
QY 1 MLFNSVLRQPLQGLVLRNGSSQYVQLSLTGYQCSGNDEHTSYGTVGVPPVPGCTFFSA 60
DB 1 MLFNSVLRQPLQGLVLRNGSSQYVQLSLTGYQCSGNDEHTSYGTVGVPPVPGCTFFSA 60
QY 61 PNMHVLAVANBEGFVRLYNTSOSPRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB 61 PNMHVLAVANBEGFVRLYNTSOSPRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
QY 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKPEKAVFCTGGRDGNIMVWDTCKKXGDFVR 180
DB 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKPEKG-----YQ 156
QY 181 QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQSVTVVLFDQENTLVSAGAVDGI 240
DB 157 -----DFQSVTVVLFDQENTLVSAGAVDGI 183
QY 241 KWDLRKNTAVRQEPFIASKGYLPGSSRKLGYSLILDSTGTLFANCTDDNIYMFNM 300
DB 184 KWDLRKNTAVRQEPFIASKGYLPGSSRKLGYSLILDSTGTLFANCTDDNIYMFNM 243
QY 301 TGLKTSFVALFNCHONSTFVKSSLSPPDQFLVSGSSDEAAYTWKYSTPQPTVLLGHS 360
DB 244 TGLKTSFVALFNCHONSTFVKSSLSPPDQFLVSGSSDEAAYTWKYSTPQPTVLLGHS 303
QY 361 QEVTSVCMCPSPDKTIATCDNTLKIWRNLGRLEKPGDKLSTVGMASQKKESRPG 420
DB 304 QEVTSVCMCPSPDKTIATCDNTLKIWRNLGRLEKPGDKLSTVGMASQKKESRPG 363
QY 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACPACAGDLPLPSNTPFTSIPAKARSPI 480
DB 364 VVTSSQSTPAKAPRVKCNPSNSSPSAACPACAGDLPLPSNTPFTSIPAKARSPI 423
QY 481 NRRGSVSVSPKPPSSFKMSIRNWTPTPSSSPPIPPASETKIMSPKALIPVSKSQ 540

DB 424 NRRGSVSVSPKPPSSFKMSIRNWTPTPSSSPPIPPASETKIMSPKALIPVSKSQ 483
QY 541 AEACSESNRVKRRDLSSCLSESVKQKCVKSCNCTVTELDGOVENLHLDJCCLAGNOEDLSK 600
DB 484 AEACSESNRVKRRDLSSCLSESVKQKCVKSCNCTVTELDGOVENLHLDJCCLAGNOEDLSK 543
QY 601 DSLGPTKSSKIEGAGTISIEPPSPISPYASVSCGTLPLRPGCGGSEMVGKENSSENK 660
DB 544 DSLGPTKSSKIEGAGTISIEPPSPISPYASVSCGTLPLRPGCGGSEMVGKENSSENK 603
QY 661 NWLLAMAARKAENPSPRSPSSQTPNSRRQSGKTLPSVTTTPSSMRKICIFYFHKRSQED 720
DB 604 NWLLAMAARKAENPSPRSPSSQTPNSRRQSGKTLPSVTTTPSSMRKICIFYFHKRSQED 663
QY 721 FCGPEHSTEL 730
DB 664 FCGPEHSTEL 673
RESULT 10
ABU09610
ID ABU09610 standard; protein; 729 AA.
XX AC ABU09610;
XX DT 01-JUL-2003 (first entry)
DE Mouse retinoic acid-regulated nuclear matrix-associated protein.
KW Nucleic acid identification; T cell receptor; T cell; tumour antigen;
KW antigen presenting cell; HLA; nucleic acid library; ELISPOT assay;
KW enzyme-linked immunospot; T cell receptor binding epitope;
KW cancer associated antigen; ramp; mouse; polytope;
KW retinoic acid-regulated nuclear matrix-associated protein.
OS Mus musculus.
XX US2003003485-A1.
XX PD 02-JAN-2003.
XX PF 14-MAY-2002; 2002US-00145396.
XX PR 15-MAY-2001; 2001US-0291125P.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PI Uenaka A, Nakayama E;
XX WPI; 2003-416593/39.
XX Identifying nucleic acids encoding epitopes that bind T cell receptors on
PT T cells when presented by human leukocyte antigen by using an ELISPOT
PT assay that detects factors secreted by T cell in response to the receptor
PT binding.
XX Example 6; Fig 13; 51pp; English.
XX The invention describes a method of identifying nucleic acid encoding an
CC epitope that binds to a T cell receptor on T cell. The method involves
CC coculturing antigen presenting cells that contain a nucleic acid library
CC and express HLA molecules presenting epitopes, with a T cell having a T
CC cell receptor that binds the epitope so that the T cell receptor binds
CC the epitope encoded by the nucleic acid library, detecting a factor
CC secreted by the T cell in response to T cell receptor binding, using
CC ELISPOT (enzyme-linked immunospot) assay, and correlating factor
CC secretion with presence of a nucleic acid encoding the epitope. The
CC method is useful for identifying a nucleic acid molecule encoding the
CC epitope that specifically binds to a T cell receptor on a T cell when
CC presented by an HLA molecule. A second method described in the invention
CC is useful for identifying antigens that specifically bind to a T cell
CC receptor on a T cell. The method is useful for identifying cancer

CC associated antigens. This is the amino acid sequence of the mouse
CC retinoic acid-regulated nuclear matrix-associated protein (ramp) peptides
CC of which can be combined with peptides of one or more other cancer-
CC associated antigens to form polytopes

XX	Sequence 729 AA;	
QY	Query Match 89.7%; Score 3479.5; DB 6; Length 729;	
DB	Best Local Similarity 89.2%; Pred. No. 1.7e-262;	
DB	Matches 651; Conservative 36; Mismatches 42; Indels 1; Gaps 1;	
QY	1 MLFNSVLRQPLGVLNRGNSQYPLQSLLTGYQCSGNDHTSYGETGVPVPPFGCTFSSA 60	
DB	1 MLFNSVLRQPLGVLNRGNSQYPLQSLLTGYQCSGNDHTSYGETGVPVPPFGCTFCTA 60	
QY	61 PNMEHLAVANEGFVRLYNTESQSFRRKCFKEMAHNAVFDLAWVPGELKLVTAGDQ 120	
DB	61 PSMEHLAVANEGFVRLYNTESQSFRRKCFKEMAHNAVFDLAWVPGELKLVTAGDQ 120	
QY	121 TAKFMDVKAGELGTCKGHQCSLKSVAFKPEKAVFCTGGGRGNIMVWDRCKNKDGFYR 180	
DB	121 TAKFMDVRAGELGTCKGHQCSLKSVAFKPEKAVFCTGGGRGNIMVWDRCKNKDGFYR 180	
QY	181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFFQSVTVVLFQDENTILVSAGVDGII 240	
DB	181 QVNQISGAHNTADKQTPSKPKKQNSKGLAPVDSQSVTVVLFQDENTILVSAGVDGII 240	
QY	241 KVDLRKNYTAHQEPIASKSFYLPQSGSTRKGLYSSILDTSGTSTLPANCTDDNIYFMNM 300	
DB	241 KVDLRKNYTAHQEPIASKSFYLPQSGSTRKGLYSSILDTSGTSTLPANCTDDNIYFNT 300	
QY	301 TGLKTSVPAIFNGHONSTFVVKSSLPDDQFLVSGSDDEAAYTWKYSTPQPTVLLGHS 360	
DB	301 TGLKTSVPAIFNGHONSTFVVKSSLPDDQFLVSGSDDEAAYTWKYSMPHPPPTVLLGHS 360	
QY	361 QEYTSVCMQPSDFTKIATCSDNTLKIWLRNLRGEEKPGDGLSTVGWASQKKESRRPGL 420	
DB	361 QEYTSVCMQPSDFTKIATCSDNTLKIWLRNLRGEEKP-GDKHSIVGWSQKKKEVKACP 419	
QY	421 VVTSSQSTPAKAPRVKCNPSNPSAACAAPCAGDLPLPSNTPFTSIKTSAPAKASPI 480	
DB	420 VTVPSQSTPAKAPRAKSSPSISSPSAACTPSCAGDLPLPSSPTFTSVKTTPTATRSSV 479	
QY	481 NREGSVSSVSPKPPSKFASIRNWTPTPSSSPITPPASETKIMSPRKALIPVSQSSQ 540	
DB	480 SRRGSISSVSPKPLSFKFMSLRNWTPTPSSSPVTPPASETKISSPRKALIPVSQSSQ 539	
QY	541 AEACSESRNRVXRLLDSSCLSEYKQKCVKSCNCTELDGOVENLHLDLCLAGNQBDLSK 600	
DB	540 ADACSESRNRVXRLLDSSCLSEYKQKCVKSCNCTELDGOAESLRLLDCLSGTQEVLSQ 599	
QY	601 DSLGPTKSSKIEGAGTSISPPSPISPYASESCGTPLPLRPCGEGSEMWGKNSPENK 660	
DB	600 DSEGPTKSSKTEGAGTSISPPSPISPYASESCGTPLPLRPCGEGSEMWGKNSPENK 659	
QY	661 NWLLAAAKKKAENPSPRSSOTPHSRROSGKTLSPVTITPSSMRKICTYFHRKSOED 720	
DB	660 NWLLAAAKKKAENPSPRSSOTPHSRROSGKTLSPGPVTITPSSMRKICTYFRRKTQDD 719	
QY	721 FCGPEHSTEL 730	
DB	720 FCSPEHSTEL 729	

RESULT 11

ADQ97694
ID ADQ97694 standard; protein; 786 AA.
XX AC ADQ97694;
XX DT 07-OCT-2004 (first entry)
XX DE Mouse cancer associated sequence MP10-029, SEQ ID 671.

XX	Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse.	
KW	Mus musculus.	
OS	WO2004060304-A2.	
PN	22-JUL-2004.	
XX	22-DEC-2003; 2003WO-US041389.	
PD	27-DEC-2002; 2002US-00330773.	
XX	(SAGR-) SAGRES DISCOVERY INC.	
XX	Morris DW, Malandro MS;	
PI	WPI; 2004-543781/52.	
DR	New isolated cancer associated nucleic acids comprising at least 10	
XX	contiguous nucleotides, useful for diagnosing, preventing and/or treating	
PT	cancers such as leukemia and lymphoma.	
PT	Claim 1; SEQ ID NO 671; 199pp; English.	
XX	The present invention relates to cancer associated sequences (ADQ97025-	
CC	ADQ98004). The sequences are useful for the diagnosis, prevention and/or	
CC	treatment of cancer, such as leukemia and lymphoma. Note: The sequence	
CC	data for this patent did not form part of the printed specification, but	
CC	was obtained in electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences.	
CC	Sequence 786 AA;	
QY	Query Match 75.1%; Score 2913.5; DB 8; Length 786;	
DB	Best Local Similarity 70.5%; Pred. No. 3.2e-218;	
DB	Matches 570; Conservative 34; Mismatches 46; Indels 159; Gaps 5;	
QY	1 MLFNSVLRQPLGVLNR-----	17
DB	39 MLFNSVLRQPLGVLNRGAPWIAADCGPGLPRPERIRASEDRRTSCLGSPRTLTFTVI 98	
QY	18 -----GWSQYPLQSLLTGYQCSGND 38	
DB	99 ELFSSERRPKGRSVASRAPGCLRQYRALTRBCTLRVCGNSHYPLOSLLSYQNCND 158	
QY	39 EHTSYGETGVPVPPFGCTFSSAPNMEHLAVANEGFVRLYNTESQSFRRKCFKEMAHW 98	
DB	159 EHTSYGETGVPVPPFGCTFCTAPSMEHLAVANEGFVRLYNTESQTSKTCFK----- 212	
QY	99 NAVFDLAWVPGELKLVTAGDQTAKFDVKAGELGTCKGHQCSLKSVAFSKPEKAVFCT 158	
DB	213 -----	212
QY	159 GGRDGNIMVWDRCKNKDGFYQVNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOOS 218	
DB	213 GERDGNIMVWDRCKNKDGFYQVNOISGAHNTADKQTPSKPKKQNSKGLAPAVDSQOS 272	
QY	219 VTVVLFQDENTILVSAGVDGIIKWMDLRKNYTAHQEPIASKSFYLPQSGSTRKGLYSSLI 278	
DB	273 VTVVLFQDENTILVSAGVDGIIKWMDLRKNYTAHQEPIASKSFYLPQSGSTRKGLYSSLV 332	
QY	279 LDSTGSTLPANCTDDNIYFMNMTGLKTSVPAIFNGHONSTFVVKSSLPDDQFLVSGSSD 338	
DB	333 LDSTGSTLPANCTDDNIYFMNMTGLKTSVPAIFNGHONSTFVVKSSLPDDQFLVSGSSD 392	
QY	339 EAAYTWKVSTPQPTVLLGHSQEVTSVCMQPSDFTKIATCSDNTLKIWLRNLRGEEKP 398	
DB	393 EAAYTWKVSNPWHPPPTVLLGHSQEVTSVCMQPSDFTKIATCSDNTLKIWLRNLRGEEKP 452	
QY	399 GGDKLSTVGWASQKKESRPGVTVVTSQSTPAKAPRVKCNPSNPSAACAAPCAGDL 458	
DB	453 -GDKHSIVGWSQKKKEVKACPVTVPSSQSTPAKAPRAKSSPSISSPSAACTPSCAGDL 511	

QY 459 PLPSNTPTTSIKTSPAKARSPINRRGSSVSSGPKPPSSPKMSIRKNVWTRTPSSSPPTTPP 518
DB 512 PLPSSTPTTSVKTTPATTTSSVSRGSSISGSPKPLSSPKMSLRNWNVTRTPSSSPPTTPP 571
QY 519 ASETKIMSPRKALIPVQSKSOAEACSESRNRVRRRLDSSCLESVKQKCVKSCNCVTELD 578
DB 572 ASETKLSSPRKALIPVQSKSQADACESRNRVRRRLDSSCLESVKQKCVKSCNCVTELD 631
QY 579 GOVENLHLDCLCLAGNQEDLSGLPTKSSKIEGAGTISBPSPSPISPYASESCGTLP 638
DB 632 GQAESLRDLCLCLSGTQEVLSQDSEPTKSSKTEGAGTISBPSPSPVYASEGCGPLPL 691
QY 639 PLRPGCEGSEMGKENSSENKNWLLAMAOKAENPSRSPSSQTPNSRRRSGKTLPLSP 698
DB 692 PLRPGCEGSEMGKENSSENKNWLLAIAAKKAENSSPRSPSSQTPNSRRRSGKTLSPG 751
QY 699 VTIPTSSMRKIC-----TYFHRKSQEDPCG 723
DB 752 STL A-----VCGNHVQLVDAGIKCEFCG 774

RESULT 12
ID AAB92779 standard; protein; 434 AA.
XX
AC AAB92779;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:11264.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000BP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 11264; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 434 AA;

Query Match 59.1%; Score 2293; DB 4; Length 434;
Best Local Similarity 100.0%; Pred. No. 4.2e-170;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 297 MFNMTGLKTSVPVAFNGHONSTFYVKSSLSPPDQFLVSGSSDEAAAYIWKVSTPWPPTVL 356
DB 1 MFNMTGLKTSVPVAFNGHONSTFYVKSSLSPPDQFLVSGSSDEAAAYIWKVSTPWPPTVL 60

QY 357 LGHSQEVTSVCWCPDFTKIATCSDDNLTAKIWRNLNRGLEEKPGDKLSTVGWASQKKES 416
DB 61 LGHSQEVTSVCWCPDFTKIATCSDDNLTAKIWRNLNRGLEEKPGDKLSTVGWASQKKES 120

QY 417 RPLVTVTSSOSTPAKAPRVKCNPSNSPSSAACAPSCAGDLPLPSNTPTTSIKTSPAKA 476
DB 121 RPLVTVTSSOSTPAKAPRVKCNPSNSPSSAACAPSCAGDLPLPSNTPTTSIKTSPAKA 180

QY 477 RSPINRGSVSVPSPKPPSSPKMSIRKNVWTRTPSSSPPTTPPASSTKIMSPRKALIPVQ 536
DB 181 RSPINRGSVSVPSPKPPSSPKMSIRKNVWTRTPSSSPPTTPPASSTKIMSPRKALIPVQ 240

QY 537 KSSQAEACSESRNRVRRRLDSSCLESVKQKCVKSCNCVTELDGQVENLHLDCLCLAGNQ 596
DB 241 KSSQAEACSESRNRVRRRLDSSCLESVKQKCVKSCNCVTELDGQVENLHLDCLCLAGNQ 300

QY 597 DLKSDSLGPTKSSKIEGAGTISBPSPSPISPYASESCGTLPPLRPGCEGSEMGKENS 656
DB 301 DLKSDSLGPTKSSKIEGAGTISBPSPSPISPYASESCGTLPPLRPGCEGSEMGKENS 360

QY 657 PENKNWLLAMAOKAENPSRSPSSQTPNSRRRSGKTLPSVTTTPSSMRKICTYFHRK 716
DB 361 PENKNWLLAMAOKAENPSRSPSSQTPNSRRRSGKTLPSVTTTPSSMRKICTYFHRK 420

QY 717 SQEDFCGPEHSTEL 730
DB 421 SQEDFCGPEHSTEL 434

RESULT 13
ADM20092
ID ADM20092 standard; protein; 332 AA.
XX
AC ADM20092;
XX
DT 20-MAY-2004 (first entry)
XX
DE Protein encoded by novel human channel/transporter gene #164 clone 2.
XX
KW immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
KW cyostatic; cardant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; opthalmological;
KW gene therapy; channel/transporter protein; rheumatoid arthritis;
KW neoplasm; cardiac arrest; cerebrovascular disorder; cerebral ischemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease;
KW ocular disorder; corneal infection; wound healing;
KW epithelial cell proliferation; skin aging; sunburn; transplantation;
KW chemotaxis; food additive.
OS Homo sapiens.
XX

PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-02556719P.

PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-476159/51.
DR N-PSDB; ADM19367.
XX
XX
PT Isolated nucleic acid molecule encoding a channel/transporter protein is
PT used in preventing, treating or ameliorating a medical condition.
XX
XX
PS Claim 11; SEQ ID NO 653; 809pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
CC channel/transporter protein or sequences at least 95% identical to a
CC these. The nucleic acids and proteins encoded by them are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. The antibodies to the proteins can also be used
CC in alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiodenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. This sequence corresponds to a protein of the
CC invention.
XX
SQ Sequence 287 AA;
Query Match 36.4%; Score 1413; DB 4; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.3e-101;
Matches 266; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 109 GELKLVTAAGDQTAKEFDVKAGELIGTCCKHOCSLKSVAFSKFEKAVFTCGRDGNIMYW 168
Db 3 GELKLVTAAGDQTAKEFDVKAGELIGTCCKHOCSLKSVAFSKFEKAVFTCGRDGNIMYW 62
Qy 169 DTRCNKKGDFYRQVNIQISGAHNNTSDKQTPSKPKKQNSKGLAPSDVDFQOSVTVLFDQEN 228
Db 63 DTRCNKKGDFYRQVNIQISGAHNNTSDKQTPSKPKKQNSKGLAPSDVDFQOSVTVLFDQEN 122
Qy 229 TLVSAGAVDGI1KKVWDLRKNYTAIRQEP1ASKSFYLPGSSTRKGLGYSSILIDSTGTFLPA 288
Db 123 TLVSAGAVDGI1KKVWDLRKNYTAIRQEP1ASKSFYLPGSSTRKGLGYSSILIDSTGTFLPA 182
Qy 289 NCTDDNIYMFNMTGLKTSVAIFNGHONSTFVYKSLSPDDQFLVSGSDDEAAYIKVST 348
Db 183 NCTDDNIYMFNMTGLKTSVAIFNGHONSTFVYKSLSPDDQFLVSGSDDEAAYIKVST 242
Qy 349 PWQPPTVLLGHQSOEVTSVCMCPDFT 374
Db 243 PWQPPTVLLGHQSOEVTSVCMCPDFT 268
RESULT 15
AAB92715

ID AAB92715 standard; protein; 232 AA.
XX AAB92715;
AC
XX
XX
DT 26-JUN-2001 (first entry)
XX
XX
DE Human protein sequence SEQ ID NO:11130.
XX
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX
OS Homo sapiens.
XX
XX
PN EP1074617-A2.
XX
XX
PD 07-FEB-2001.
XX
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
XX
PR 29-JUL-1999; 99JP-00248036.
XX
PR 27-AUG-1999; 99JP-00300253.
XX
PR 11-JAN-2000; 2000JP-00118776.
XX
PR 02-MAY-2000; 2000JP-00183767.
XX
PR 09-JUN-2000; 2000JP-00241899.
XX
XX
PA (HELI-) HELIX RES INST.
XX
XX
PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
PI Tehii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX
DR WPI; 2001-318749/34.
XX
XX
PT Primer sets for synthesising polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX
PS Claim 8; SEQ ID NO 11130; 2537pp + Sequence Listing; English.
XX
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 232 AA;

Query Match 31.4%; Score 1216; DB 4; Length 232;
Best Local Similarity 99.6%; Pred. No. 2.4e-86;
Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 499 MSIRNVWTRTPSSSPITPPASTKIMSPRKALIPVSQSSQAEACSESRNRVKRLDSS 558
DB 1 MSIRNVWTRTPSSSPITPPASTKIMSPRKALIPVSQSSQAEACSESRNRVKRLDSS 60
QY 559 CLESVKQKCVKSCNVCVTDLGQVHHLDLCLCLAGNQEDLSKDSLGLPTKSSKIEGAGTSI 618
CLESVKQKCVKSCNVCVTDLGQVHHLDLCLCLAGNQEDLSKDSLGLPTKSSKIEGAGTSI

Db 61 CLESVKQKCVKSCNVCVTDLGQVHHLDLCLCLAGNQEDLSKDSLGLPTKSSKIEGAGTSI 120
QY 619 SEPPSPISPYASESCGTLPLPLRPGCGEGSEMVGKENSSENKWNWLLAAAKRAENPSPR 678
DB 121 SEPPSPISPYASESCGTLPLPLRPGCGEGSEMVGKENSSENKWNWLLAAAKRAENPSPR 180
QY 679 SPSSQTPNSRRQSGKTLPSVTTITPSSMRKICTYFHRKSQEDFCGPEHSTEL 730
DB 181 SPSSQTPNSRRQSGKTLPSVTTITPSSMRKICTYFHRKSQEDFCGPEHSTEL 232

Search completed: November 8, 2006, 18:08:35
Job time : 204 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 05:53:24 ; Search time 13047 Seconds
(without alignments)
12133.642 Million cell updates/sec

Title: US-10-726-160-1
Perfect score: 2831
Sequence: 1 ggcacgagcgagggttgag.....atgaaaaaaaaaaaaaa 2831

Scoring table: IDENTITY NUC
Gapop 10,0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est3:*
 - 3: gb_est4:*
 - 4: gb_est5:*
 - 5: gb_est6:*
 - 6: gb_htc:*
 - 7: gb_est2:*
 - 8: gb_est7:*
 - 9: gb_est8:*
 - 10: gb_est9:*
 - 11: gb_gse1:*
 - 12: gb_gse2:*
 - 13: gb_gse3:*
 - 14: gb_gse4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1814	64.1	4205	6	AK166351	Mus muscu
2	1813.2	64.0	4180	6	AK133177	Mus muscu
3	1802.2	63.7	2350	6	AK054412	Mus muscu
4	1800.6	63.6	4179	6	AK164249	Mus muscu
5	1777.8	62.8	4201	6	AK169601	Mus muscu
6	1758	62.1	4228	6	AK154594	Mus muscu
7	1627.4	57.5	1923	14	DQ043568	Homo sapi
8	1539.2	54.4	3655	6	AK161401	Mus muscu
9	1451.8	51.3	1923	14	DQ043569	Pan trogl
10	996.2	35.2	1106	1	AL557830	AL557830
11	992.8	35.1	998	1	AL553783	AL553783
12	976.4	34.5	1042	1	AL576790	AL576790
13	927.2	32.8	1075	4	BX416133	BX416133
14	890	31.4	4308	6	AK164999	Mus muscu
15	885.8	31.3	949	3	BQ431493	AGENCOURT
16	881.6	31.1	1014	3	BM800170	AGENCOURT
17	874.6	30.9	907	4	BX462391	BX462391
18	856.4	30.3	905	1	AU123749	AU123749
19	852.2	30.1	958	1	AL558302	AL558302

20	851.4	30.1	1128	2	BM458369	AGENCOURT
21	847.4	29.9	914	3	BUI176525	AGENCOURT
22	845.8	29.9	962	1	AL580551	AL580551
23	845	29.8	929	3	BUI508613	AGENCOURT
24	824.2	29.1	899	4	CA488488	AGENCOURT
25	820.4	29.0	922	1	AL558817	AL558817
26	816.8	28.9	924	10	DV918737	DV918737
27	804.4	28.4	853	1	AU124656	AU124656
28	799.4	28.2	883	3	BQ233905	AGENCOURT
29	794	28.0	902	3	BQ34727	AGENCOURT
30	788.2	27.8	866	4	EX435948	EX435948
31	786	27.8	804	8	CX164038	AGENCOURT
32	760.6	26.9	820	5	CD651902	CD651902
33	755	26.7	936	2	BI257561	BI257561
34	751.6	26.5	1012	3	BM557720	AGENCOURT
35	749	26.5	843	9	DA736257	DA736257
36	733.4	25.9	1107	2	EM468859	AGENCOURT
37	731	25.8	731	9	CX786033	CX786033
38	729.4	25.8	902	3	BU521670	AGENCOURT
39	725	25.6	727	8	CN362845	CN362845
40	719.2	25.4	847	8	CV556506	CV556506
41	712.2	25.2	962	7	BE746556	BE746556
42	708	25.0	708	8	CN362847	CN362847
43	707.4	25.0	763	2	BG17017	BG17017
44	701.4	24.8	981	7	BE742123	BE742123
45	700.8	24.8	711	9	DN996329	DN996329

ALIGNMENTS

RESULT 1	AK166351	4205 bp	mRNA	linear	HTC 21-SEP-2005
LOCUS	Mus musculus mammary gland RCB-0526 Jyg-MC(A) cDNA, RIKEN full-length enriched library, clone:G830048A21 product:L2D7L protein, full insert sequence.				
DEFINITION	AK166351				
ACCESSION	AK166351.1	GI:74147704			
VERSION	HTC; CAP trapper.				
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.				
REFERENCE	1	Carninci, P. and Hayashizaki, Y.			
AUTHORS	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	10349636				
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
AUTHORS	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
TITLE	Genome Res. 10 (10), 1617-1630 (2000)				
JOURNAL	11042159				
PUBMED	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
AUTHORS	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipicapillary sequencer				
TITLE	Genome Res. 10 (11), 1757-1771 (2000)				
JOURNAL	11076861				
PUBMED	4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,			

Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Momberte, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

11217851

CONSRMTM
JOURNAL
PUBMED
REFERENCE
AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusich, V., Chochia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L. M., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reid, J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

FANTOM Consortium

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420 (6915), 563-573 (2002)

12466851

CONSRMTM
JOURNAL
PUBMED
REFERENCE
AUTHORS

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V. B., Bremner, S. E., Batalov, S., Forrest, A. R., Zavolan, M., Davis, M. J., Wilming, L. G., Aidinis, V., Allen, J. E., Ambesi-Impombato, A., Apweiler, R., Aturaliya, R. N., Bailey, T. L., Bansal, M., Baxter, L., Belsel, K. W., Bersano, T., Bono, H., Chalk, A. M., Chiu, K. P., Choudhary, V., Christoffels, A., Clutterbuck, D. R., Crowe, M. L., Dalla, E., Dalrymple, B. P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C. F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T. R., Gojobori, T., Green, R. E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T. K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanehisa, A., Katoh, M., Kawasawa, Y., Kelsso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S. P., Kruger, A., Kummerfeld, S. K., Kurochkin, I. V., Lareau, L. F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S.,

Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Parg, K. C., Pavan, W. J., Pavese, G., Peole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J. F., Ring, B. Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S. L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C. A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugura, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S. L., Tang, S., Taylor, M. S., Tegner, J., Teichmann, S. A., Ueda, H., van Nimwegen, E., Verardo, R., Wei, C. L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimond, S. M., Teasdale, R. D., Liu, E. T., Brusich, V., Quackenbush, J., Wahlestedt, C., Mattick, J. S., Hume, D. A., Kal, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, T., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Sano, H., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.

FANTOM Consortium

The transcriptional landscape of the mammalian genome
Science 309 (5740), 1559-1563 (2005)

16141072

7

CONSRMTM
JOURNAL
PUBMED
REFERENCE
AUTHORS

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakaniishi, M., Nakamura, M., Nishida, H., Tap, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batalov, S., Engstrom, P. G., Mizuno, Y., Faghihi, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.

RIKEN Genome Exploration Research Group

Antisense transcription in the mammalian transcriptome
Science 309 (5740), 1564-1566 (2005)

16141073

8 (bases 1 to 4205)

CONSRMTM
JOURNAL
PUBMED
REFERENCE
AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CONSRMTM
JOURNAL
PUBMED
REFERENCE
AUTHORS

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

1. 4205
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="FANTOM DB:G830048A21"
/db_xref="taxon:10090"
/clone="G830048A21"
/cell_line="RCB-0526 JYG-MC(A)"
/tissue_type="mammary gland"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
123. 2312
/note="unnamed protein product; L2DTL protein
(UniProt|Q8BW38, evidence: FASTA, 99.7%ID, 100%length, match=2187)
putative"

CDS

```

/codon_start=1
/protein_id="BAE38725.1"
/db_xref="GI:74147705"
/translation="MLFNSVLQRPQLGVLRNGSSHYPLQSLLSGVQCNDBHSTYSY
ETGVVPPFGTCTAPSMHEHLAVANEEGFVRLYNTESTQTSKTKCFKEMAHNAVF
DLAWFVGLKLTAAQDQTAKEWDRAGELMTCKHQCSLKSFAFKQAVFSTGG
RDCNIMIDWTRCNKKGDFYRQVQNGISGAHNTAKPTPKPKKQNSKGLAPAVDSQOS
VTVLFGDNTLVSACAVDGIILKMDLRKNYTAIROEPIASKSLFLYPGTSTRKLGVSS
LVLDSTGLFLANCTDDNLYMENMTGLKTSVAVENGHQNSTFYVYKSSLSPPDQPLIS
GASDEAAYLWKVSMHPHPPTVLLHGSQEVTSVCMSPDFIKIATCSDNLTILKWLNR
GUEEKDSDHISVGMTSKKKEVKACPTVPSSQSTPAKAPRAKSSPSSSSAACT
PSCAGDLPTSPSTPTSVKTTTATTRSSVRSRGSISVSPKLSLPSFQMLRNWVTRP
SSSPPTVPASPTKSSPKALIPVQSKSQADACSESRNRVKRLRDLSDCLSVKQK
VKSNCVTELDQGBSLRLDCLSLGAEQVLSQDSBPTKSSKTEGAGTSSISEPPSPV

```

Query Match		64.1%;	Score 1814;	DB 6;	Length 4205;
Best Local Similarity		86.9%;	Pred. No. 0;		
Matches 2018;		Conservative 0;	Mismatches 300;	Indels 4;	Gaps 2;
QY	8	GGGGAGTTGGAGCGGATAACGATTGTGTGTGAGAGCGCAACGTCGGATTCTCTGCTG	67		
DB	6	GGGGAGTTGGAGCGCAACCGCGCGGTGTGTGGAGGTTGCAGACGGCGGCTCTGTGG	65		
QY	68	AACCTCGAGGC-ATTTCACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATG	126		
DB	66	AACCTGGGGCGATCTCTCGGGGTTCCCTCAGCTGAGCCTTCTCCTCAGGCCCAGATG	125		
QY	127	CTCTTCAATTCGTGCTCGCCAGCCAGCTGTTGGCGTCTGAGAAATGATGGTCTTCA	186		
DB	126	CTCTTCAACTCGGTGCTCGCCAGCCGAGCTCGGGTCTCTGAGGAACGGGTGCTTCA	185		
QY	187	CAATACCTCTTCAATCCCTTCTGACTGTGTATCAGTCAGTGAGTGAATGATGAACACT	246		
DB	186	CATTACCTCTCAATCCCTTCTTAAGTGGTATCAGTGCACACTGTAAACGATGAACACG	245		
QY	247	TCTTATGGAGAAA CAGGAGTCCAGTTCCTCTCTTTTGGATGTACTTCTCTCTGCTCCC	306		
DB	246	TCTTATGGAGAAACAGGAGTCCAGATTCTCTCTTTTGGATGACCTTCTGTACCGCTCCC	305		
QY	307	AATATGGAACTGACTAGCAGTTGCCAATGAAGAGGCTTTGTCGATTGTATTAACACA	366		
DB	306	AGTATGGAGCATATATTAGCAGTTGCTAATGAAGAGGCTTTGTCAGATTATATATACA	365		
QY	367	GAATCAAAAGTTTCAGAAAGAGTCTTCAAGAAATGGATGGCTCACTGGAAATGCCGTC	426		
DB	366	GAGTCAAAATAGCAAAAAGACATGCTTCAAGAGTGGATGGCTCACTGGAAATGCTGTC	425		
QY	427	TTTGACTTGGCTGGGTCTCTGGTGAACTTAAACTTTGTTACAGCAGCAGGTGATCAAA	486		
DB	426	TTTGACTTGGCTGGGTCTCTGGTGAACTTAAACTTTGTTACAGCAGCAGGTGATCAAA	485		
QY	487	GCCAAATTTTGGAGCTAAAGCTGTGTGAGCTGATTGGAACTGCAAGAGTCAATGTC	546		
DB	486	GCCAAATTTTGGAGCTGAGAGCTGTGTGAGCTAATGGAACTGCAAGAGGCTCAAGTGC	545		
QY	547	AGCCTCAAGTCAGTTCCTTTCTAAGTTTGAGAAAGCTGATTCTGTACGGGTGGAAG	606		
DB	546	AGCCTCAAGTCAGTTCCTTTCTAAGTTTCAAAAGCTGTGTTCTTACAGGGGGGAGA	605		
QY	607	GATGGCAACATTTATGTCTGGGATACAGGTGCAACAAAAGATGGGTTTTATAGGCAA	666		
DB	606	GACGGCAACATTTATGATCTGGGACACAGGTGTAAACAAAAGATGATTTATAGACAA	665		
QY	667	GTGAATCAATCAGTGGAGCTCAATACCTCAGACAAGCAACCCCTTCAAAACCCAAAG	726		
DB	666	GTGAATCAATCAGTGGAGCTCAATACCTCAGACAAGCAACCCCTTCAAAAGCCCAAG	725		
QY	727	AGAAACAGATTTCAAAAGACTTGTCTCTTCTGTGATTTCCAGCAAAAGTGTACTGTG	786		
DB	726	AGAAACAAATTTCAAAAGACTTGTCTCTGCTGTGGATTCCAGCAGAGTGTACTGTG	785		
QY	787	GTCTCTTTTCAAGCAGAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAA	846		

DB	786	GTCTCTTTTTCAGGATGAGATAATAGTCTCAGCAGGAGCGGTGGATGAATAATCAAA	845		
QY	847	GTATGGGATTTACGTAAGAAATTTATCTCTTATCACAAGAACCCATAGCATCCCAAGTCT	906		
DB	846	GTATGGGATTTTCGCAAGAAATTTACACTGCTTATCACAAGAACCCATAGCATCCCAAGTCT	905		
QY	907	TTCTCTGTATCCCAAGGTAGCAGCACTCGAAACCTTGGATATTCAAGTCTGATTTTGGATTC	966		
DB	906	TTCTCTGTATCCCAAGGTATCCAGCACTCGAAAGCTAGGATACTCGAGTTTGGTTTAGACTCT	965		
QY	967	ACTGGCTCTACTTTTATTTGCTTAATTTGCACAGACGATTAACATCTACATGTTTAAATGACT	1026		
DB	966	ACTGGCTCTACTTTTATTTGCTTAATTTGCACAGATGACAACATCTATATGTTTAAATGACT	1025		
QY	1027	GGGTTTGAAGACTTCTCCAGTGGCTATTTTCAATGAGCACCAAGACTCTACCTTTTATGTA	1086		
DB	1026	GGCTTAAAGACTTCTCCGGTGGCTGTCTTCAATGAGCACCAAGACTCTACCTTTTATGTA	1085		
QY	1087	AAATCCAGCTTGTAGTCCAGATGACCAAGTTTTGTAGTCAAGTGGCTCAAGTGAAGCTGCC	1146		
DB	1086	AAATCAAGTCTTGTAGTCCAGATGACCAAGTTTTGTAGTCAAGTGAAGCTGCC	1145		
QY	1147	TACATATGAAGGCTCTCCACACCTGGGAACTCTCTACTGTGCTCTGGGTCAITCTCAA	1206		
DB	1146	TACATTTGGAAGGTTTCCATGCGCATCCCTCTACTGTGCTCTGGGTCAITCTCAA	1205		
QY	1207	GAGGTCAAGTCTGTGCTGTGCTGCTCATCTGACTTTCACAAAGATTGCTACCTTGTCTGAT	1266		
DB	1206	GAGGTCAAGTCTGTGCTGTGCTGCTCATCTGACTTTCACAAAGATTGCTACCTTGTCTGAT	1265		
QY	1267	GACAAATACACTAAAAATCTGGCGCTTGAATGAGGCTTAGAGGAGAAAAACAGAGGTGAT	1326		
DB	1266	GATAATACACTGAAATCTGGCGCTTGAATGAGGCTTAGAGGAGAAAAAC--AGGTGAT	1322		
QY	1327	AACTTTTCCACGGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACTCGGCTTAGTA	1386		
DB	1323	AAACATTTCCATAGTGGGTTGGACCTCTCAGAAAGAAAAAGAGTCAAGGCTTCCCAAGTA	1382		
QY	1387	ACAGTAAACGAGTACGAGTACTCTGCCAAGCCCCCAGGGTAAAGTCAATCCATCC	1446		
DB	1383	ACGGTACCAAGTACGAGTACTCTGCCAAGCTTCCAGAGCTCCAGAGCAGTCCATCC	1442		
QY	1447	AACTTCTTCCCCGTCACTCGCAGCTTGTGCCCCCAAGCTGTCTGGAGACCTCCTCTTCTCT	1506		
DB	1443	ATCTCTCTCTCTCTCTCTCAGCAGCTTGTACTCCGAGCTGTGCAGGAGACCTCCTCTTCTCT	1502		
QY	1507	TCAAACTACTCTTACCTTCTTATTAACCTCTCTGCCAAGCCCGGCTCTCCCATCAAC	1566		
DB	1503	TCAAGTACCCCACTTCTCAGTCAAAACCACTCTCTGCCACGACCCGTTCTTTCAGTCAAC	1562		
QY	1567	AGAAAGGCTCTGCTCTCTCTCTCTCTCCAGCCACCTTCACTCTTCAAGATGTCGATT	1626		
DB	1563	AGAAAGGCTCTCTCTCTCTCTCTCTCTCTCCAGCCACCTTCACTCTTCAAGATGTCGATT	1622		
QY	1627	AGAAACTGGGTGACCCGAAACACTTCTCTCATCACCAACCCATCACTCCACTCTCTGGGAG	1686		
DB	1623	AGAAACTGGGTGACCCGAAACACTTCTCTCATCACCAACCTGTCTCTCTCTCTCTGAG	1682		
QY	1687	ACCAAGATCATGCTCTCCGAGAAAAGCCCTTATTCCTGTGAGCCAGAGTCAATCCCAAGCA	1746		
DB	1683	ACAAAGATCTCATCTCTCAAGAAAAGCTTATTTCTGTGAGCCAGAAAGTCAACAGGCA	1742		
QY	1747	GAGGCTTGTCTGAGCTTAGAAATAGAGTAAAGAGGCTAGACTCAAGCTGTCTGGAG	1806		
DB	1743	GATGCTTGTCTGAAATCTAGAAATAGAGTAAAGAGGCTTGTGACTCAAGCTGTCTGGAG	1802		
QY	1807	AGTGTGAAACAAAGTGTGTGAAGTGTGAATCTGTGACTGTGACTTGTATGGCCCAAGTT	1866		
DB	1803	AGTGTGAAACAAAGTGTGTGAAGTGTGAATCTGTGACTGTGACTTGTATGGCCCAAGG	1862		
QY	1867	GAAAACTTTTCAATTTGGATCTGTGCTGCCCTTGTGTGTAACACAGGAGACCTTGTATAGGAC	1926		
DB	1863	GAGAGTCTTGTGATCTGTGCTGCCCTTCCGGCGCCAGGAGTCTTCTTAGCCCAAGAC	1922		

Qy	1927	TCCTAGTCCCTACCAATCAAGCAAAATTAAGGAGCTGGTACCAGTATCTCAGAGCCT	1986
Db	1923	TCGAGGGTCTTACCAATCAAGCAAGACTGAAGGTGGCAAGCATCTCAGAACCT	1982
Qy	1987	CGTCTCTTATCAGTCCGATGCTTCAGAAAGCTGTGGAACCTACCTCTTCTCTTTGAGA	2046
Db	1983	CTTCTCTCTGTCAGTCTTATGCTTCTGAAGGCTGTGGACCACTGCTCTTCTCTTTGAGA	2042
Qy	2047	CTTGTGGAGAAAGGGTCTGAATGTGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAGC	2106
Db	2043	CTTGTGGAGAAAGGATCTGAGATGTGTGGCAAGAGAAATAGTCTCCAGAGAAATAAAGC	2102
Qy	2107	TGTTGTGGCCATGCGACGCAAAAGCAAGGCTGAGATCCATCTCCAGAAAGTCCGCTCA	2166
Db	2103	TGTTGTGGCCATGAGCAGCAAAAGCAAGGCAAGAAATTCATCCCAAGAAAGTCCATCA	2162
Qy	2167	TCCCAGACACCAATTCAGGAGAGACAGCGCAAGACATTCGCAAGCCGCTCACCATC	2226
Db	2163	TCTCAGACACCCAGTTCAGGAGAGAAAGTGGAAGACGTCACCAAGCCGCTCACCATT	2222
Qy	2227	AGCCAGCTCCATGAGGAAATCTGCACATACTTCCATAGAAAGTCCCGAGGAGCTTC	2286
Db	2223	ACTCCAGCTCCATGAGGAAATGTACATACTTCTGTAGAAGACTCAAGATGACTTC	2282
Qy	2287	TGTGTCCTGAACACTCAACAGAAATATAGATCTTAATCTGA	2328
Db	2283	TGCAGTCTTGAACACTCAACTGAATATATAGATGCTAATCTGA	2324
RESULT 2			
AK133177			
LOCUS	AK133177	4180 bp mRNA linear	HTC 21-SEP-2005
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4931420L14 product:l2DTL protein, full insert sequence.		
ACCESSION	AK133177		
VERSION	AK133177.1	GI:74221710	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
1			
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res.	10 (10), 1617-1630	(2000)
PUBMED	11042159		
2			
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res.	10 (11), 1757-1771	(2000)
PUBMED	11076861		
3			
REFERENCE	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,		

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikolaou, I., Pesole, G., Schombach, C., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.	
RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium	
FUNCTIONAL annotation of a full-length mouse cDNA collection	
Nature 409 (5821), 585-590 (2001)	
11217851	
5	
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikolaou, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusci, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kasawara, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Mikki, H., Nagashima, T., Numa, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempie, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.	
FANTOM Consortium	
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
Nature 420 (6915), 563-573 (2002)	
12466851	
6	
Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakachi, H.,	

REFERENCE AUTHORS	5 The FANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	6 RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the FANTOM Consortium.
TITLE	Antisense Transcription in the Mammalian Transcriptome
JOURNAL REFERENCE	Science 309, 1564-1566 (2005)
AUTHORS	7 The FANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).
TITLE	The Transcriptional Landscape of the Mammalian Genome
JOURNAL REFERENCE	Science 309, 1559-1563 (2005)
AUTHORS	8 (bases 1 to 2350) Adachi, J., Alizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayaishi, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Maesuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayaishi, K., Y.
TITLE JOURNAL	Direct Submission
AUTHORS	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers 1. 2350 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:E330023F01" /db_xref="taxon:10090" /clone="E330023F01" /sex="female" /tissue types="ovary" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="2 days pregnant adult" 105..2294 /notes="unnamed protein product; L2DTL PROTEIN (RA-REGULATED NUCLEAR MATRIX-ASSOCIATED PROTEIN) homolog [Homo sapiens] (SPTR Q9NZJ0, evidence: FASTA, 89.5%ID, 100%length, match=2187) putative" /codon_start=1 /protein_id="BAC35769.1" /db_xref="GI:26344225"
CDs	translations="MLPNSVLRQPLQGLRWGSSHYPLQSLISGYQCNDBHTSYG ETCVPPFGCTCTAPSEHILAVANESGFVRLYNTEQSQTCKTCFKEWMAHNAVE DLAWVGEKLVTAGDQKAFKWDVRAGELMGCKHQCSLSKSVAFPKQKAVFTGG RDGNIMWDRNCNKGDFYRQVNIQISGAHNTADKQTPSPKPKQNSKGLAPVSDQS VTVLFDQNTLVSAAGDGIKVDLRKNYTAQROEPIASKSFLYPTGSTRKLGYS LVLDSGTSLFANCTDNTIMFNMTGLKTSFPAVFNHQNSTFYVKSLSLSDDDQLIS GSSDEAYIWKVSMWPHPTVLLGHGSEVTSVCWCFSDFTKIATCSDDDNTLKIWLNR

|||||
903 ACCAGACTCGAAGCTAGGATACCTCGAGTTGGTTTGTAGACTCTACTGGCTCTACTTTA 962
982 TTTTGTAAATTCACAGACCATTAACATCTACATGTTTAAATATGACTGGGTTGAAGACTTCT 1041
963 TTTTGTAACTGCACAGATGACAAATCTATATGTTCAATATGACTGGCTTAAAGACTTCT 1022
1042 CCAGTGGCTATTTTCAATGGAACACAGAACTCTACCTTTTATGTTAAATTCAGCCTTAGT 1101
1023 CCGGTGGCTGTCTTCAATGGAACACAGAACTCTACCTTTTATGTTAAATTCAGTCTTAGT 1082
1102 CCAGATGACAGTTTATGCTAGTGGCTCAAGTGATGAAGCTGCTTACATATGGAAGGTC 1161
1083 CCAGATGACAGTTTATGCTAGTGGCTCAAGTGATGAAGCTGCTTACATTTGGAAGGTT 1142
1162 TCCACACCTCGGCAACCTCTACTGTCTTCTGGGTCAATTTCAAGAGGTCACGTCTGTG 1221
1143 TCCATGCCATGGCATCTCTACTGTCTTCTGGGTCAATTTCAAGAGGTCACGTCTGTG 1202
1222 TCGTGGTTCATCTGATCTTCAAAAGATGCTACCTGTTCCTGATGACAAATACACTAAA 1281
1203 TCGTGGTTCATCTGATCTTCAAAAGATGCTACCTGTTCCTGATGATTAATACACTGAAA 1262
1282 ATCTGGCGCTTGAATAGAGCTTAGAGGAGAAACAGGAGGTGATAACTTTCCAGGTTG 1341
1263 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAAC---AGGTGATAAAACATTTCCATAGTG 1319
1342 GGTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGCTAGTACAGTAACAGAGTAGC 1401
1320 GGTGGACCTCTCAGAGAAAAGAGAGTGAAGGCTGCCAGTAAAGGTCACCAAGTAGC 1379
1402 CAGAGTACTCTGCCAAAGCCCGAGGTAAGTGCAATTCATTCCTTTCCCGGTCA 1461
1380 CAGAGTACTCTGCCAAAGCTCCAGAGCAAGAGCAGTCCATCCATCTCTCTCTCTCG 1439
1462 TCGGAGCTGTGGCCCAAGCTGTGCTGAGACCTCCCTCTCTCTTCAATTAATCTCTAG 1521
1440 TCAGCAGCTGTGTACCTCGAGCTGTGAGAGACCTCCCTCTCTCTTCAAGTACCCCA 1499
1522 TTCTCTATTAAACCTCTCTGCGAAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTG 1581
1500 TTCTCAGTCAAAACCACTCTGCGCAGACCCGGTTCTTCTCAGTCAGCAGAGAGGCTCCATC 1559
1582 TCTCCGCTCTCTCCCAAGCACTTTCATCTTTCAAGATGTCATAGAACTGGGTGACC 1641
1560 TCTTCTGTCTCTCCCAAGCACTCTCATCTTTCAAGATGTCGCTTAGAACTGGGTGACC 1619
1642 CGAACACCTTCTCATCACCACCCATCACTCCACCTGCTTCGGAGACCAAGATCATGCT 1701
1620 CGAACACCTTCTCTCATCAACCTGTCTCTCAGCTGCTTCGAGCAAAAGATCTCATCT 1679
1702 CCGAGAAAAGCCCTTATCTCTGTGACCCAGAGTCACTCCCAAGCAGAGGCTTGTCTGAG 1761
1680 CCAAGAAAAGCTTTATCTCTGTGACCCAGAGTCACTACAGCAGATGCTTGTCTGAA 1739
1762 TCTAGAAATAGAGTAAAGAGGCTTAGACTCAAGCTGTCTGAGAGTGTGAAAACAAAG 1821
1740 TCTAGAAATAGAGTAAAGAGGCTTGTGACTCAAGCTGTCTGAGAGTGTGAAAACAAAG 1799
1822 TGTGTGAAGAGTTGTAATGTTGACTGTGAGCTGTGAGCCAGTTGAAAATCTTCAATTGG 1881
1800 TGTGTGAAGAGTTGCAACTGTGTCTGTGCTGTGAGCTGTGAGCCAGCGGAGAGTCTTGTG 1859
1882 GATCTGTGCTGCTCTGCTGTAACCAAGGAGACCTTAGTAAGGACTCTCTAGTCTCTACC 1941
1860 GATCTGTGCTGCTTTCGGGCAACCAAGAGTCTCTTAGCCAAGACTCCGAGGCTCTACC 1919
1942 AAATCAAGCAAAATTAAGAGGCTGTAACAGTATCTCAGAGCTCTCGTCTCTCTATCACT 2001
1920 AAATCAAGCAAGACTGAAGGTGTGGCAACAGCATCTCAGAACTCTCTCTCTCTGCTGCT 1979
2002 CCGTATGCTTCAGAAAGCTGTGAAACGCTACTCTCTCTCTTGTGAGACCTTGTGGAGAGG 2061
|||||

Db 1980 CCTATGCTTCTGAAGGCTGTGGACCACTGCTCTTCTCTTGTGAGACCTTGTGGAGAGGA 2039
Qy 2062 TCTGAATGTGTAGCAAGAAATAGTTTCCCAGAGAAATAAAACTGTTGTGGCCATG 2121
Db 2040 TCTGAGATGTGTGGGCAAGAAATAGCTTCTCCAGAGAAATAAGAACTGTTGTGGCCATA 2099
Qy 2122 GCAGCCAAACGGAAGGCTGGAATCCATCTCCAGAGTCCGTCATCCAGACACCCCAAT 2181
Db 2100 GCAGCCAAACGGAAGGCAAGAAATTTTATCTCCCAAGAAATCCATCATCTCAGACACCCAGT 2159
Qy 2182 TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCCACCATCACGCCAGCTCCATG 2241
Db 2160 TCCAGGAGACAAAGTGGGAAGACGTCACCAAGCCCGGTCCACCATTTACTCTCCAGCTCCATG 2219
Qy 2242 AGGAAATCTGCACATATCTCCATAGAAAAGTCCCAAGGAGGACTTCTGTGGTCTCTGAAAC 2301
Db 2220 AGGAAGATATGTACATATCTTTCGTAGAAAAGACTCAAGATGACTTCTGCACTCTCTGAAAC 2279
Qy 2302 TCAACAGAAATTATAGATTCTTAATCTCGA 2328
Db 2280 TCAACTGAATTATAGATGCTAATCTCGA 2306

RESULT 4
AK164249 4179 bp mRNA linear HTC 21-SEP-2005
DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:D030041C04 product:l2DPL protein, full
insert sequence.

ACCESSION AK164249
VERSION AK164249.1 GI:74211282
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Tozawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staab, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,

Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashina, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
11217851

5

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, I., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kwasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Negashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

FANTOM Consortium

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420 (6915), 563-573 (2002)
12466851

6

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Implombato, A., Apweiler, R., Attalaya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhury, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Faghihi, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashizaki, Y., Hensch, T.K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelson, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugtara, K., Sultana, R.,

Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlstedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanishi, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.

FANTOM Consortium

The transcriptional landscape of the mammalian genome
Science 309 (5740), 1559-1563 (2005)
16141072

7

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yagi, C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlstedt, C.

RIKEN Genome Exploration Research Group

Antisense transcription in the mammalian transcriptome
Science 309 (5740), 1564-1566 (2005)
16141073

8 (bases 1 to 4179)

Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan [E-mail: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

1. 4179
/organism="Mus musculus"
/mol_type="mRNA"
/strains="C57BL/6J"
/db_xref="FANTOM DB:D030041C04"
/db_xref="taxon:10090"
/clone="D030041C04"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="9 days embryo"

misc_feature
1. 4179
/note="L2DTL protein (UniProt|Q8BW38, evidence: FASTY, 99.9%ID, 100%length, match=2188)
putative"

ORIGIN

Query Match 63.6%; Score 1800.6; DB 6; Length 4179;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 2015; Conservative 0; Mismatches 299; Indels 5; Gaps 3;

Qy 12 GAGTTGGAGCGCATACGATTTGTGTGTGTGACAGCGCAACGTGCGATTTCTGCTGAAT 71
|||||
|||||

1 GAGTTGGAGCGAAACCGCGCGGTTGTGGAGGGTGACAGCGCGGCTCTCTGTGGAACT 60
72 TGGAGGC-ATTTCTAGCACTTTTCTCTCAGCTGAGGCTTTTCTCGGACCCCTGATGCTCT 130
61 TGGGGGCGATCTCTCGGCGTTCCCTCAGCTGAGCGCTTCTCCTCAGGCCCCAGATGCTCT 120
131 TCAATTCGGTGTCTCGGCCAGCCCCAGCTTGGCGTCTCTGAGAAATGGATGGTCTTCACAAAT 190
121 TCAACTCGGTACTCCGCGAGCGAGCTCGGCGTCTCTGAGGAACGGGTGGTCTTTCACATT 180
191 ACCCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACACACTTCTT 250
181 ACCCTCTCAATCCCTTCTAAGTGGTTATCAGTGCAACTGTAACGATGAACACACAGTCTT 240
251 ATGGAGAAACAGGAGTCCAGTTCCTCTTTTGGATGTACCTTCTCTCTGCTGCCAATA 310
241 ATGGAGAAACAGGAGTCCAGTTCCTCTTTTGGATGCACCTTCTGTACCGCTCCCGTA 300
311 TGGAAACATGTACTAGCAGTTGCCAATGGAAGAGGCTTTGTTGCGATTGTATATACAGAAAT 370
301 TGGAGCATATATTAGCAGTTGCTAATGAAGAAGGCTTTGTGAGATTATATATAACAGAAAT 360
371 CACAAAGTTTCAGAAAGAGTCTTCACAAAGAAATGGATGGCTCACTGGAAATGCCGCTTTTG 430
361 CACAAACTAGCAAAAAGACATGCTTCAAGGAGTGGATGSGCTCACTGGAAATGCTGCTTTG 420
431 ACTGGCTTGGGTCTCTGTGTAACCTTAACTTTGTTACAGCAGCAGGTGATCAACAGGCA 490
421 ACTTGGCTTGGGTCTCTGTGTAACCTTAACTTTGTTACAGCAGCAGGTGATCAACAGGCA 480
491 AATTTTGGGACCTGAAAAGCTGTGTGAGCTGATTGGAAACATGCCAAAGGTCAATCAATGCAGC 550
481 AATTTTGGGATGTAGAGCTGTGTGAGCTGATGGGACATGCAAAAGGCCACCAAGTGCAGCC 540
551 TCAA-GTCAAGTTGCCCTTTTCTAAGTTTGAAGAAGCTGTATTCTGTACGGGTGGAAGAGAT 609
541 TCAAGTCTGTAGCTTTTCCCGAGTTTCAAAAAGCTGTGTTCTCTACAGGGGGAGAGAC 600
610 GGCACATTATGGTCTGGGATACGAGGTGCCAACAAAAAGATGGGTTTTATAGGCAAGTG 669
601 GGCACATTATGATCTGGGACACCGAGTGTAAACAAAAAGATGGATTTTATAGACAAAGTG 660
670 AATCAAAATCAGTGGAGCTCACAAATACCTCAGACACAGCAAAACCCCTTCAAAACCCCAAGAG 729
661 AATCAAAATCAGTGGAGCTCACAAATACCTCAGACACAGCAAAACCCCTTCAAAACCCCAAGAG 720
730 AAACAGAAATTCAAAAGGACTTGTCTCTCTGTGGATTTTCAGCAAAAGTGTACTGTGGTC 789
721 AAACAGAAATTCAAAAGGACTTGTCTCTCTGTGGATTTCCACAGCAGAGTGTACTGTGGTC 780
790 CTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGAGGCTGTGGATGGGATATCAAGTA 849
781 CTCTTTCAGATGAGAATAATATAGTCTCAGCAGGAGCGGTGGATGGAATAATCAAGTA 840
850 TGGGATTTACGTAAGAAATATACGCTTATCGCAAGACCCATAGCATCCCAAGTCTTTC 909
841 TGGGATTTTCGCAAGAAATATACGCTTATCGCAAGAACCCATAGCATCCCAAGTCTTTC 900
910 CTGTACCCAGGTAGCAGACCTCGAAAACCTTGGATATTCAAGTCTGATTTTGGATTCCACT 969
901 CTGTACCCAGGTACCGACCTCGAAGCTAGGATCTCGAGTTGGTTTGTAGCTCTACT 960
970 GGCCTACTTTTATTTGCTAAATTTGCACAGACGATAACATCTACATGTTTAAATATGATGGG 1029
961 GGCCTACTTTTATTTGCTAACTGACAGATGACACATCTATATGTTTCAATATGATGGC 1020
1030 TTGAAGACTTCTCCAGTGGCTATTTTCAATGACACCGAACTCTACCTTTTATGTAAAA 1089
1021 TTAAAGACTTCTCCGGTGGCTGTCTTCAATGACACCGAACTCTACCTTTTATGTAAAA 1080
1090 TCCAGCTTTAGTCCAGATACCAAGTTTATTTAGTCAGTGGCTCAAGTGAAGCTGCCTAC 1149
1081 TCAAGTCTTTAGTCCAGATGACCAGTTTAAATCAGTGGTTTCAAGTGTGAAGCTGCCTAC 1140

1150 ATATGGAAGGTCTCACACCCTGGCAACCTCTCTACTGTGCTCTGGGTCAATTTCTCAAGAG 1209
1141 ATTTGGAAAGTTTCAATGCCATGGATCTCTCTACTGTGCTCTGGGTCAATTTCTCAAGAG 1200
1210 GTCACGTCTGTGCTGTGTGCTTCATCTGAACTTCAACAAAGATTGCTACTGTTCTGATGAC 1269
1201 GTCACGTCTGTGCTGTGTGCTTCATCAGACTTCCAAAGATTGCAACCTGCTCTGATGAT 1260
1270 AATACACTAAAAATCTGGCGCTTGAAATAGAGGCTTAGAGGAGAAACCAAGGAGGTGATAAA 1329
1261 AATACACTGAAAAATCTGGCGCTTGAAATAGAGGCTTAGAGGAGAAACCC--AGGTGATAAA 1317
1330 CTTTCCACGGTGGGTGGGCTCTCAGAGAAAGAAAGAGTCAAGACCTGGCTCTAGTAACA 1389
1318 CATTCATAGTGGGTGGACCTCTCAGAAAGAAAAGAAAGTGAAGACCTGCCAGTAACG 1377
1390 GTAAAGAGTAGCCAGAGTACTCTGCCCCAAAGCCCCCAGGGGTAAAGTGCAAATCCATCCAAT 1449
1378 GTACCAAGTAGCCAGAGTACTCTGSCCAAAGCTCCCAGAGCCNAGAGCAGTCCATCCATC 1437
1450 TCTTCCCCGTCACTCGCAGCTTGTGCCCCAAAGCTGTGCTGGAGACCTCCCTTCTCTTCA 1509
1438 TCTCTCTCTCTGTCAGCAGCTTGTACTCCGAGCTGTGAGGAGACCTCCCTCTCTCTTCA 1497
1510 AATACTCTAGTTCCTCTATTAAACCTCTCTGSCCAAAGGCCGTCTCCCATCAACAGA 1569
1498 AGTACCCCCCACTTCTCAGTCAAAACCACTCTGCGCAGACCCGTTCTTCTAGTCAGCAGA 1557
1570 AGAGCTCTGTCTCTCTCGCTCTCTCCAAAGCACCTTCTCATCTTTCAAGATGTCGATTAGA 1629
1558 AGAGCTCCATCTCTTCTGTGCTCTCCAAAGCCACTCTCATCTTTCAAGATGTCGCTTAGA 1617
1630 AACTGGGTGACCCGAAACCTTCTCTCATCAACCCATCATCTCCACTCTGCTTTCGGAGACC 1689
1618 AACTGGGTGACCCGAAACCTTCTCTCATCACACCTGTCTCACTCCCACTGCTTCTGAGACA 1677
1690 AAGATCATGTCTCCGAGAAAGCCCTTATTCTCTGTGAGCCAGAGTCAATCCCAAGCAGAG 1749
1678 AAGATCTCATCTCCAAAGAAAAGCTCTTATTCTGTGAGCCAGAGTCAATCAGGCAAGAT 1737
1750 GCTTGTCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGAGAGAT 1809
1738 GCTTGTCTCTGAATCTAGAAATAGAGTGAAGGCGCTCTTGACTCAAGCTGTCTGAGAGT 1797
1810 GTGAAAAAAAGTGTGGAAGAGTGTAACTGTGTGACTGAGCTTGTGATGGCCAAAGTTGAA 1869
1798 GTGAAAAAAAGTGTGGAAGAGTGTCAACTGTGTCTAGCTTGTGACGCGCCAAAGCGGAG 1857
1870 AATCTTCAATTTGGATCTGTGCTGCTTGTCTGTAACCAAGGAAGACCTTAGTAAGGACTCT 1929
1858 AGTCTTCTGTTGGATCTGTGCTGCTTTCGCGCAACCCAGGAAGTCTCTTAGCCAAAGCTCC 1917
1930 CTAGTCTCTACCAATCAAGCAAAATTCGAAGGAGCTGGTACAGTATCTCAGAGGCTCCG 1989
1918 GAGGTCTTACCAATCAAGCAAGCTGAAGGTGTGCAACAGCATCTCAGAACTCTCT 1977
1990 TCTCTTATCAGTCTCGTATGCTTTCAGAAAGCTGTGGAACGCTACCTCTCTCTCTTGTAGACT 2049
1978 TCTCTGTCTGCTCTTATGCTTTCGAAGGCTGTGACCACTGCTCTCTCTCTCTTGTAGACT 2037
2050 TGTGGAGAGGCTCTGAAATGTTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACCTGG 2109
2038 TGTGGAGAGGATCTGAGATGTTGGGCAAGAGAAATAGCTCTCCAGAGAAATAAGAACTGG 2097
2110 TTGTTGGCCATAGGACGCAAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCCGCTATCC 2169
2098 TTGTTGGCCATAGGACGCAAAACGGAAGGCAAGAAATTCATCCCAAGAAAGTCCATCACT 2157
2170 CAGACACCAATTTCCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGTCCACCTCAGG 2229
2158 CAGACACCGTTCCAGAGACAAAGTGGGAAGAGCTCACCAGGGCCCCGTCCACCTACT 2217

QY 2230 CCAGCTCCATGAGGAAAATCTGCACATACCTTCCATAGAAAGTCCAGAGGACTTCTGT 2289
|||||
Db 2218 CCAGCTCCATGAGGAGATATGTACATCTTCTGTAGAAAGACTCAAGTACACTTCTGC 2277
|||||
QY 2290 GGTCTGTAACACTCAACAGAAATTTATGATCTTAATCTGA 2328
|||||
Db 2278 AGTCCTGAACACTCAACTGAATTTATGATGCTAATCTGA 2316
|||||

RESULT 5
AK169601
LOCUS
DEFINITION
Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN
full-length enriched library, clone:E430001J12 product:L2DTL
protein, full insert sequence.
AK169601
AK169601.1 GI:74143874
HTC; CAP trapper.
KEYWORDS
Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
High efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Taahiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Harai, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., Kling, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J.S., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Botunga, N.,
Carninci, P., de Bona, M.F., Brownstein, M.J., Bult, C.,
Fleischmann, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohtsuki, S.
and Hayashizaki, Y.
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
11217851
5

AUTHORS
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J.S., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempke, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Harai, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
FANTOM Consortium
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420 (6915), 563-573 (2002)
12466851
6
Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C.,
Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R.,
Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R.,
Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E.,
Ambesi-Impombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L.,
Bansal, M.P., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M.,
Chlu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R.,
Crome, M.D., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G.,
di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G.,
Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M.,
Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E.,
Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N.,
Hill, D., Huminecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T.,
Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Keiso, J., Kitamura, H.,
Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K.,
Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J.,
Liu, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L.,
Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S.,
Morris, K., Mortazavi-Tabar, S., Mulder, N., Nakano, N., Nakachi, H.,
Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O.,
Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G.,
Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z.,
Ringwald, C., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A.,
Schneider, C., Schonbach, C., Sekiguchi, K., Sempke, C.A., Seno, S.,
Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D.,
Sincclair, B., Sperling, S., Stupka, E., Sugita, K., Sultana, R.,
Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S.,
Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R.,
Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A.,
Hilde, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T.,
Bruic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A.,
Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M.,
Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M.,
Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M.,
Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M.,
Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K.,
Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.
FANTOM Consortium
The transcriptional landscape of the mammalian genome
Science 309 (5740), 1559-1563 (2005)

Gustincich, S., Hirokawa, N., Jackson, J.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kwasawa, Y., Kedierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, K., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wellie, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, K., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

FANTOM Consortium
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420 (6915), 563-573 (2002)
12466851

6
Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Bremner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Imponente, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Banaei, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Dingemans, P., Englstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Humnietz, L., Iacono, M., Ikeo, K., Iwano, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kele, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuoka, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Moris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Sempke, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiu, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.

FANTOM Consortium
The transcriptional landscape of the mammalian genome
Science 309 (5740), 1559-1563 (2005)
16141072

7
Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Englstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.

CONSRMT
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

RIKEN Genome Exploration Research Group
Antisense transcription in the mammalian transcriptome
Science 309 (5740), 1564-1566 (2005)
16141073

8 (bases 1 to 4228)
Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watanahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

COMMENT

Location/Qualifiers
1. 4228
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM_DB:F630050P04"
/db_xref="taxon:10090"
/clone="F630050P04"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
1. 4228
/note="L2DTL protein (UniProt|Q8BW38, evidence: FASTY, 98.5%ID, 100%length, match=2221)
putative"

FEATURES
source

misc_feature
1. 4228
/note="L2DTL protein (UniProt|Q8BW38, evidence: FASTY, 98.5%ID, 100%length, match=2221)
putative"

ORIGIN

Query Match 62.1%; Score 1758; DB 6; Length 4228;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 2023; Conservative 0; Mismatches 295; Indels 38; Gaps 5;
QY 8 GCGGAGTTGGAGCGGATACGATTTGTGTGTGAGAGCGGCGACGTCGATTTCTGTG 67
6 GCGGAGTTGGAGCGGAGGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 65
QY 68 AACTTGGAGGC-ATTTCTACGACTTTTCTCAGCTGAGGCTTTTCTCCGACCCCTGATG 126
66 AACTTGGGCGGCGATTTCTCGCGCGTTCCTCCTCAGCTGAGCGCTTCTCCTCAGCGCCAGATG 125
QY 127 CTCCTTCAATTCGGTCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 186
126 CTCCTTCAATTCGGTCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 185
QY 187 CAATACCCCTCTTCAATCCCTTCTGACTGTTATCAGTGCAGTGGTGAATGATGAACACACT 246
186 CATTACCCCTCTACATCCCTTCTTAAGTGGTTATCAGTGCATGTAACGATGAACACACG 245
QY 247 TCTTATGAGAACAGGAGTCCAGTTCCTCTTTTGGATGTACCTTCTCTTCTCTCTCC 306
246 TCTTATGAGAACAGGAGTCCAGTTCCTCTTTTGGATGTACCTTCTCTTACCCTCTCC 305
QY 307 AATATGGAACTGTACTAGCAGTTCCCAATGAAGAGCGCTTGTTCGATTTGATTAACACA 366

KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
REFERENCE 1 (bases 1 to 1923)
AUTHORS Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeello, D., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees
JOURNAL (ez) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 1923)
AUTHORS Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeello, D., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.
FEATURES Location/Qualifiers
source 1..1923
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
gene <1..>1923
/gene="RAMP"
/locus_tag="HC10244"
ORIGIN
Query Match 57.5%; Score 1627.4; DB 14; Length 1923;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 1812; Conservative 0; Mismatches 111; Indels 171; Gaps 1;
QY 124 ATGCTCTTCAATTCGGTGTCTCGCCAGAGCCCGAGCTTGGCGTCTGAGAAATGGATGGTCT 183
DB 1 ATGCTCTTCAATTCGGTGTCTCGCCAGAGCCCGAGCTTGGCGTCTGNGAAATGGATGGTCT 60
QY 184 TCACAATACCTTCAATCCCTTCTGACTGTGATATCAGTGCAGTGGTAAATGATGAACAC 243
DB 61 TCACAATACCTTCAATCCCTTCTGACTGTGATATCAGTGCAGTGGTAAATGATGAACAC 120
QY 244 ACTTCTTATGGAGAACAGAGTCCCGAGTTCCTCTTTTGGATGTACCTTCTCTCTGCT 303
DB 121 ACTTCTTATGGAGAACAGAGTCCCGAGTTCCTCTTTTGGATGTACCTTCTCTCTGNN 180
QY 304 CCCAATATGGAAACATCTACTAGCAGTTGCCAATGAAGAGGCTTTGTCGATGTATAC 363
DB 181 NNN 240
QY 364 ACAGAAATCAAAAGTTTCAGAAAGAGTGTCTCAAGAAATGGATGGCTCACTGGAAATGCC 423
DB 241 NNN 300
QY 424 GTCTTTGACTGGCTGGCTTCTGCTGAATTAATTAATTTGTTACAGCAGCAGGTGATCAA 483
DB 301 GTCTTTGACTGGCTGGCTTCTGCTGAATTAATTAATTTGTTACAGCAGCAGGTGATCAA 360
QY 484 ACAGCCAAATTTTGGACGTAAAGCTGGTGAAGCTGATTTGGAACATGCAAGAGTCAATCAA 543
DB 361 ACAGCCAAATTTTGGACGTAAAGCTGGTGAAGCTGATTTGGAACATGCAAGAGTCAATCAA 420
QY 544 TGCAGCCTCAAGTCAGTTCCTTTTCTAAGTTTGAAGAAAGCTGTATTTGTACGGGTGGA 603
DB 421 TGCAGCCTCAAGTCAGTTCCTTTTCTAAGTTTGAAGAAAGCTGTATTTGTACGGGTGGA 460
QY 604 AGAGATGGCAACATTATGTCTGGGATACAGGTGGCAACAAAGAAAGATGGGTTTATAGG 663

461 ----- 460
664 CAAGTGAATCAATCAGTGGAGCTCAATATACCTCAGAAAGCAAAACCCCTTCAAAACCC 723
461 ----- 460
724 AAGAAGAAACAGAATTCAAAGGACTTGCTCTTCTGTGGATTTCACGAAAGTGTACT 783
461 -----GCTATCAGGATTTCCAGCAAGTGTACT 489
784 GTGGTCTCTTTCAAGACGAGAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATAATC 843
490 GTGGTCTCTTTCAAGACGAGAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATAATC 549
844 AAGATATGGGATTTACGTAAAGATTAATCTGCTTATCGACAAGAACCCATAGCATCCAAAG 903
550 AAGATATGGGATTTACGTAAAGATTAATCTGCTTATCGACAAGAACCCATAGCATCCAAAG 609
904 TCTTTCTGTACCCAGGTAGCAGCAGCTCGAAACCTTGGATATTCAAGTCTGATTTTGGAT 963
610 TCTTTCTGTACCCAGGTAGCAGCAGCTCGAAACCTTGGATATTCAAGTCTGATTTTGGAT 669
964 TCCACTGGCTCTACTTTATTTGCTAAATTCACAGACGATACATCTACATGTTTAATATG 1023
670 TCCACTGGCTCTACTTTATTTGCTAAATTCACAGACGATACATCTACATGTTTAATATG 729
1024 ACTGGTTGAAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTAT 1083
730 ACTGGTTGAAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTAT 789
1084 GTAAATCCAGCCTTAGTCCAGATGACGAGTTTTTGTAGTCAAGTGGCTCAAGTGAAGACT 1143
790 GTAAATCCAGCCTTAGTCCAGATGACGAGTTTTTGTAGTCAAGTGGCTCAAGTGAAGACT 849
1144 GCCTACATATGAAGGCTCTCCACACCCCTGGCAACCTCTACTGTGCTCCTGGGTCATTCT 1203
850 GCCTACATATGAAGGCTCTCCACACCCCTGGCAACCTCTACTGTGCTCCTGGGTCATTCT 909
1204 CAAGAGGTACGTCTGTGCTGTGCTGTGCTCAATCTGACTTCAAAAAGATTTGCTACCTGTTCT 1263
910 CAAGAGGTACGTCTGTGCTGTGCTGTGCTCAATCTGACTTCAAAAAGATTTGCTACCTGTTCT 969
1264 GATGCAATATCACTAAATCTGGCGTTGAAATAGAGGCTTAGAGGAGAAACCCAGGAGGT 1323
970 GATGCAATATCACTAAATCTGGCGTTGAAATAGAGGCTTAGAGGAGAAACCCAGGAGGT 1029
1324 GATAAACTTTCCAGCGTGGGTTGGGCTCTCAGAGAAAGAAAGAGTCAAGACCTGGCCCTA 1383
1030 GATAAACTTTCCAGCGTGGGTTGGGCTCTCAGAGAAAGAAAGAGTCAAGACCTGGCCCTA 1089
1384 GTAAACAGTAAAGAGTAGCAGTACTCTCTGCAAAAGCCCGCCAGGGTAAAGTGAATCA 1443
1090 GTAAACAGTAAAGAGTAGCAGTACTCTCTGCAAAAGCCCGCCAGGGTAAAGTGAATCA 1149
1444 TCCAAATCTTCCCGTCAATCCGAGCTTGTGCCCAAGCTGTGCTGGAGACTCCTCCTCTT 1503
1150 TCCAAATCTTCCCGTCAATCCGAGCTTGTGCCCAAGCTGTGCTGGAGACTCCTCCTCTT 1209
1504 CTTTCAAAATCTCTTACGTTCTCTATTAAACCTCTCTCTGCAAGCCCGGCTCTCCCATC 1563
1210 CTTTCAAAATCTCTTACGTTCTCTATTAAACCTCTCTCTGCAAGCCCGGCTCTCCCATC 1269
1564 AACAGAAAGAGCTCTGTCT 1623
1270 AACAGAAAGAGCTCTGTCT 1329
1624 ATTAGAAACTGGGTGACCCGAGCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1683
1330 ATTAGAAACTGGGTGACCCGAGCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1389
1684 GAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCCAA 1743
1390 GAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCCAA 1449

Kurochkin, I. V., Lareau, L. F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K. C., Pavan, W. J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J. F., Ring, B. Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S. L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Sempie, C. A., Seno, S., Sesca, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiyama, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S. L., Tang, S., Taylor, M. S., Tegner, J., Teichmann, S. A., Ueda, H. R., van Nimwegen, E., Verardo, R., Wei, C. L., Yagi, K., Yamaniishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hilde, W., Bult, C., Grimmond, S. M., Teasdale, R. D., Liu, E. T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J. S., Hume, D. A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondou, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y. FANTOM Consortium
The transcriptional landscape of the mammalian genome
Science 309 (5740), 1559-1563 (2005)
16141072
7
Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batalov, S., Engstrom, P. G., Mizuno, Y., Feghini, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.
RIKEN Genome Exploration Research Group
Antisense transcription in the mammalian transcriptome
Science 309 (5740), 1564-1566 (2005)
16141073
8 (bases 1 to 3655)
Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (5-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. .3655
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:4930546N03"
/db_xref="taxon:10090"
/clone="4930546N03"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

112. .1905
/note="unnamed prtein product: L2DTL protein homolog [Mus musculus] (UniProt|Q8BW38, evidence: FASTY, 99.8%ID, 86.9%length, matct=1902)
putative"
/codon_start=1
/protein_id="BAE316376.1"
/db_xref="GI:74149450"
/translation="MGTCCKGHCQSLKSVAPPKFVKAVFTSGGRDNIMWDRCNKGD
AVGFIQVNOISGAHNTADQTPSPKPKKNSKGLAPVDSQSVTVVLFQDENTLVASG
ADVDIIKVDLRKNYTAIROEPIASKSPFLYPTSTRLKGLSYSLDSTGLFVNCST
DNIYMNWNTGLKTSVAVFNGHNSFTVKSLSPPDQFLISGSDDAAAYTWKSWMPW
HPPTVLLGHSEVTSVCWCPSPDFTKIATCSDDNTLKIWRNLNRGLKEPKDGHVIGWT
SOKKEVYKACVTPVPSQSPTPAKAPRAKSSPSSISPSAACTPSCAGDLPPLPSPPTIS
SVKTPATPSSVSRRGSISSVPKLSFSSFKMSLNRWVTRTPSSPPVTPPSPSTKIS
SPRKALIPVSKQSQAACSESRNRVRLDSSCLVESVKQKSCNCSNCTVLPDGOAES
LRLDCLCLSGTQEVLSQSEGPSTKSTEGAGTSISBPSPSPVSYASEGCGPLPLPLR

Query Match 54.4%; Score 1539.2; DB 6; Length 3655;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1689; Conservative 0; Mismatches 228; Indels 3; Gaps 1;
QY 409 GCTCACTGGAATGGCGTCTTTTGACCTGGCCTGGGTTCCTGGTGAACCTTAAACTTTGTTACA 468
DB 1 GCTCACTGGAATGGCGTCTTTTGACCTGGCCTGGGTTCCTGGTGAACCTTAAACTTTGTTACA 60
QY 469 GCAGCAGTGATCAAAAAGCCAAATTTTGGGACGTAAAGCTGGTGAGCTGATTTGGAACA 528
DB 61 GCAGCGGTGATCGACAGCGCAAAATTTTGGGATGTAGAGCTGGTGAGCTGATGGGGACA 120
QY 529 TGCAGAGTGCATCAATGAGCTCAAGTCAGTGGCTTTCTTAAGTTTGGAGAAAGCTGTA 588
DB 121 TGCAGAGGCCACCACTGAGCTCAAGTCAGTGGCTTTCCAGTTTCAAAAGCTGTG 180
QY 589 TTCTGTACGGTGGAGAGATGGCAACATATATGTTCTGGGATACAGGTGCAACAAAAA 648
DB 181 TTCTCTACAGGGGGAGAGACGGCAACATATATGATCTGGGACACCCAGGTGTAAACAAAAA 240
QY 649 GATGGGTTTATAGCAAGTGAATCAAAATCAGTGAGGCTCAATACCTCTCAGCAAGCAA 708
DB 241 GATGGATTTATAGCAAGTGAATCAAAATCAGTGAGGCTCAATATCTCGACAGAGGCC 300
QY 709 ACCCTTCAAAACCCCAAGAAAGACAGAAATTCAGAGGCTTGGCTTCTGTGATTTTC 768
DB 301 ACCCTTCAAAACCCCAAGAAAGACAGAAATTCAGAGGCTTGGCTTCTGTGATTTTC 360
QY 769 CAGCAAGTGTACTGTGGTCTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGAGCT 828
DB 361 CAGCAGAGTGTACTGTGGTCTCTCTTTTCAGGATGAGATACATTTAGTCTCAGCAGAGGCC 420
QY 829 GTGGATGGGATAATCAAAAGTATGGATTTAGTGAAGATTTATATCTGCTTTCGACAGAA 888
DB 421 GTGGATGGAATAATCAAAAGTATGGGATTTGGCGAAGATTTACACTGCTTATCGAAGAA 480
QY 889 CCCATAGCATCCAGTCTTCTCTGCTACCCAGGTAGCAGCACTCGAAAACTTGGATTTCA 948
DB 481 CCCATAGCATCCAGTCTTCTCTGCTACCCAGGTAGCAGCACTCGAAAACTTGGATTTCA 540
QY 949 AGTCTGATTTTGGATTTCCACTGGCTCTACTTTATTTGCTAAATTCACAGAGATTAACATC 1008
DB 541 AGTTTGGTTTATAGACTCTACTGGCTCTACTTTATTTGCTAAATTCACAGAGATTAACATC 600
QY 1009 TACATGTTTATATAGCTGGGTGAGACTTCTCAGTGGCTATTTTCAATGGACACCAAG 1068
DB 601 TATATGTTTCAATATGACTGGCTTAAAGACTTCTCCGGTGGCTGCTTCAATGGACACCAAG 660
QY 1069 AACTCTACTCTTTATGTAAAAATCCAGCTTAGTCAGATGACCACTTTTATGTAGTCACTGGC 1128
DB 661 AACTCTACTCTTTATGTAAAAATCAAGTCTTAGTCAGATGACCACTTTTATGTAGTCACTGGT 720
QY 1129 TCAAGTGTAGAGCTGCCCTACATATGGAAGGTCTCCACACCCCTGGCAACCTCTACTGTG 1188
DB 721 TCAAGTGTAGAGCTGCCCTACATATGGAAGGTCTCCATGCGCATGCGCATCTCTACTGTG 780

Qy	1189	CTCTGGGTCAATTCTCAAGAGGTCA	CGTCTGTGTGCTGGTCCATCTGACTTCA	CAAAG	1248	
Db	781	CTCTGGGTCAATTCTCAAGAGGTCA	CGTCTGTGTGCTGGTCCATCTGACTTCA	CAAAG	840	
Qy	1249	ATTGCTACCTGTTCTGATGACAAT	TACTAAACTCTGGCGCTTGAATAGAGGCT	TAGAG	1308	
Db	841	ATTGCAACCTGCTCTGATGATAAT	TACTGAAATCTGGCGCTTGAATAGAGGCT	TAGAG	900	
Qy	1309	GAGAAACGAGGAGTGATAACTTTCA	CGGTGGGTTGGCGCTCTCAGAAAGAAAAGAG	1368		
Db	901	GAGAAACG--AGTGATAAACATTC	CACTAGTGGGTTGGACCTCTCAGAAAGAAAAGAA	957		
Qy	1369	TCAAGACCTGGCTTAGTAACAGTAA	CCAGTAGCCAGAGTACTCCTGCCAAAGCCCCCAGG	1428		
Db	958	GTGAAAGCTGCTCCAGTAACGGTAC	CAAGTAGCCAGAGTACTCCTGCCAAAGCTCCCGA	1017		
Qy	1429	GTAAGTGCAATCCATCCAATTTCT	CCCCGCTCATCCGACGTTGTGCCCAAGCTGTGCT	1488		
Db	1018	GCCAAGAGCAGTCCATCCATCTCT	CTCTCTCTTCGTCAGCAGCTTGTACTCCGAGCTGTGCA	1077		
Qy	1489	GGAGACCTCCCTCTCTCTCAAA	TACTCTACGTTCTCTATTA	AAAACTCTCTCTGCCCAAG	1548	
Db	1078	GGAGACCTCCCTCTCTCTCAAG	TACCCACATCTCTAGTCA	AAACCACTCTCTGCCACG	1137	
Qy	1549	GCCCGGTCTCCCATCAACAGAA	GAGGCTGTGTCTCTCCGCTCTCTCCCAAGCCACCTTCA	1608		
Db	1138	ACCGTTCTTCTCAGTCAGCAGA	GAGGCTCCATCTCTCTGTGCTCTCCCAAGCCACTCTCA	1197		
Qy	1609	TCCTTCAAGATGTCGATAGAA	ACTGGGTGACCGGAACA	CACTTCTCTCATCA	CCACCATC	1668
Db	1198	TCCTTCAAGATGTCGCTTAGAA	ACTGGGTGACCGGAACA	CACTTCTCTCATCA	CCACCATGTC	1257
Qy	1669	ACTCCACTGCTTCGGAGACA	AGATCATGTCTCCGAGAAAGCCCTTATCTCTGTGAC	1728		
Db	1258	ACTCCA	CTGTCTTGAGACAAAGATCTCATCTCCAAGAAAGCTTTATCTCTGTGAC	1317		
Qy	1729	CAGAAGTCATCCCAAGCAGAG	GCGTTGCTCTGAGTCTAGAAATAGGTAAAGAGGAGGCTA	1788		
Db	1318	CAGAAGTCATCAGGCGAGATG	CTTCTCTGATCTAGAAATAGGTGAAGAGGCGTCTT	1377		
Qy	1789	GACTCAAGCTGCTGGAGAG	GTGAAACAAAAGTGTGTAAGAGTGTGTAATCTGTGTGACT	1848		
Db	1378	GACTCAAGCTGCTGGAGAG	GTGAAACAAAAGTGTGTAAGAGTGTGTAATCTGTGTGACT	1437		
Qy	1849	GAGCTTGATGGCCAAAGTTGA	AAATCTTTCAATTTGGATCTGTGCTGCTGCTGCTGTGTGTGTA	1908		
Db	1438	GAGCTTGACGGCCAAAGCGA	GAGTCTTCTGTTGGATCTGTGCTGCTGCTTTCGGCACCCAG	1497		
Qy	1909	GAAGACCTTAGTAAGGACTCT	TAGTCTTACCAAAATCAAGCAAAATTAAGAGGAGTGGT	1968		
Db	1498	GAAGTCTTAGCCAAAGACT	CGAGGGTCTTACCAAAATCAAGCAAGACTGAAGGTGTGGC	1557		
Qy	1969	ACCAAGTATCTCAGAGCTCC	GTCTCTCATAGTCCGATCTTCTCAGAAAGCTGTGGAACG	2028		
Db	1558	ACAAGCATCTCAGAACCTC	CTCTCTCTCTGTCAGTCTCTTATGCTTCTGAAGGCTGTGACCA	1617		
Qy	2029	CTACCTCTTCTTGGACCTT	GTGGAGAGGGTCTGAAATGTTAGGCCAAAGAGAAATAGT	2088		
Db	1618	CTGCTCTTCTTGGACCTT	GTGGAGAGGATCTGAGATGTTGGGCAAGAGAAATAGC	1677		
Qy	2089	TCCCCAGAGAAATAAACTGG	TGTTGGCCATGCGACCAAGCGAGGCTGAGAAATCCA	2148		
Db	1678	TCTCCAGAGAAATAAGAACT	GGTTGTGGCCATAGCAGC	CAAGCTGGAAGAGCGTCA	1737	
Qy	2149	TCTCCAGAAAGTCCGTCAT	CCCCAGACCCAAATTC	CAGGAGACAGAGCGGAAAGACAT	2208	
Db	1738	TCCCCAAGAAAGTCCATCAT	CTCAGACACCCAGTTCCAGGAGACAAAGTGGGAGAGCGTCA	1797		
Qy	2209	CCAAAGCCGGTCAACCATCA	CGCCAGCTCCATGAGGAAATCTTGCA	CATCTTCCATAG	2268	
Db	1798	CCAGGCCCGGTCAACCATTA	CTCCAGAGCTCCATGAGGAAGATATGTA	CATCTTCTGTAG	1857	

Qy	2269	AAGTCCAGGAGGACTTCTGTGCTCTGAACACTCAACAGAAATATAGATTTCAATCTGA	2328
Db	1858	AAGACTCAAGATGACTTCTGAGTCTCTGAACACTCAACTGAATATATAGATGCTAAATCTGA	1917
RESULT 9			
DQ043569			
LOCUS			
DEFINITION			
Pan troglodytes RAMP gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.			
ACCESSION			
DQ043569			
VERSION			
DQ043569.1 GI:66894784			
KEYWORDS			
GSS.			
SOURCE			
ORGANISM			
Pan troglodytes (chimpanzee)			
Pan troglodytes			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.			
REFERENCE			
1 (bases 1 to 1923)			
Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.			
A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees			
(er) PLOS Biol. 3 (6), E170 (2005)			
JOURNAL			
PUBMED			
15869325			
REFERENCE			
2 (bases 1 to 1923)			
Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.			
Direct Submission			
Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA			
JOURNAL			
COMMENT			
This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.			
FEATURES			
source			
1..1923			
/organism="Pan troglodytes"			
/mol_type="genomic DNA"			
/db_xref="taxon:9598"			
<1..>1923			
/gene="RAMP"			
/locus_tag="HC10244"			
ORIGIN			
Query Match 51.3%; Score 1451.8; DB 14; Length 1923;			
Best Local Similarity 78.3%; Pred. No. 0;			
Matches 1640; Conservative 0; Mismatches 283; Indels 171; Gaps 1;			
Qy	124	ATGCTCTTCAATTGCGGTGCTCCGCCAGCCCCAGCTTGGCGTCTCTGAGAAATGATGGTCT	183
Db	1	ATGCTCTTCAATTGCGGTGCTCCGCCAGCCCCAGCTTGGCGTCTCTGAGANNNGATGGTCT	60
Qy	184	TCACAATACCCCTCTTCAATCCCTTCTGACTGGTTATCAGTCGAGTGGTAAATGATGAACAC	243
Db	61	TCACAATACCCCTCTTCAATGGCTTCTGACTGGTTATCAGTCGAGTGGTAAATGATGAACAC	120
Qy	244	ACTTCTTATGGAGAAACAGGAGTCCAGTTCCTCTTTTGGATGTACCTTCTCTTCTGCT	303
Db	121	ACTTCTTATGGAGAAACAGGAGTCCAGTTCCTCTTTTGGATGTACCTTCTCTTCTGNN	180
Qy	304	CCCAATATGAAACATGTACTAGCAGTTGCCAATGAAGAAGCGCTTCTTCGATTGTATAAC	363
Db	181	NN	240
Qy	364	ACAGAATCAAAAGTTTCAGAAAGAAAGTGCTTCAAAAGAAATGATGGCTCACTGGAATGCC	423
Db	241	NN	300
Qy	424	GTCTTTGACCTGCGCTGGGTCTCTGTGTAACTTTAACTGTTACAGCAGCAGGTGATCAA	483
Db	301	GTCTTTGACCTGCGCTGGGTCTCTGTGTAACTTTAACTGTTACGSCAGCAGGTGATCAA	360

Qy	484	ACGCCAAATTTTGGACGCTAAAAAGCTGGTGAGCTGATTGGAAACATGCCAAAGGTCATCAA	543
Db	361	ACGCCAAATTTTGGACGCTAAAAAGCTGGTGAGCTGATTGGAAACATGCCAAAGGTCATCAA	420
Qy	544	TGCAGCCTCAAGTCAGTTCAGTTCCTTTTCTAAAGTTTCAGAAAGCTGTATTCTGTACGGGTGGA	603
Db	421	TGCAGCCTCAAGTCAGTTCAGTTCCTTTTCTAAAGTTTCAGAAAG-----	460
Qy	604	AGAGATGCGACACATTATGGTCTGGGATACCAGGTGCAACAAAAAGATGGTTTTATAGG	663
Db	461	-----	460
Qy	664	CAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTTCAGCAAGCAAAACCCCTTCAAAACCC	723
Db	461	-----	460
Qy	724	AAGAAGAAACAGAAATTCAAAAGGACTTGCTCCTTCTGTGGATTTCAGCAAAAGTGTACT	783
Db	461	-----GCTATCAGGATTTCCAGCAAAAGTGTACT	489
Qy	784	GTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATAATC	843
Db	490	GTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATAATC	549
Qy	844	AAAGTATGGGATTTACGTAAAGAAATTATATCTGCTTTATCGACAAGAACCCATAGCATCCAAG	903
Db	550	AAAGTATGGGATTTACGTAAAGAAATTATCTGCTTTATCGACAAGAACCCATAGCATCCAAG	609
Qy	904	TCTTTCTGTACCCAGGTAGCAGCACTCGAAAACTTCGATATTCAGATCTGATTTTGGAT	963
Db	610	TCTTTCTGTATCCAGGTAGCAGCACTCGAAAACTTCGATATTCAGATCTGATTTTGGAT	669
Qy	964	TCCACTGGCTCTACTTTTATTGCTAAATTGACACAGCATAACTACATCTGTTTAAATATG	1023
Db	670	TCCACTGGCTCTACTTTTATTGCTAAATTGACACAGCATAACTACATCTGTTTAAATATG	729
Qy	1024	ACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTAT	1083
Db	730	ACGGGTTGAAGACTTCTCCAGNNGCTATTTTCANTGGAACCAANNCTCTACCTTTTAT	789
Qy	1084	GTAATAATCCAGCCTTAGTCCAGATGACCAAGTTTTTAGTCAGTGCTCAAGTGATGAAGCT	1143
Db	790	GTAATAATCCAGCCTTTNNNNNATGACCAAGTTTTTANNNTGTGCTCAAGTGATGAAGCT	849
Qy	1144	GCCTACATATGGAAGGCTCTCACACCCTGGCAACCTCTACTGTGCTCTCTGGGTCAATTCT	1203
Db	850	GCCTACATATGGAAGGCTCTCACACCCTGGCAACCTCTACTGTGCTCTCTGGGTCAATTCT	909
Qy	1204	CAAGAGTCAAGTCTGTGTGTCTGTGTGCCATCTGACTTTCACAAAGATGCTACTGTTCT	1263
Db	910	CAAGAGTCAAGTCTGTGTGTGTGTGCCATCTGACTTTCACAAAGATGCTACTGTTCT	969
Qy	1264	GATGACAAATACACTAAATAATCTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACCAAGAGGT	1323
Db	970	GATGACAAATACACTAAATAATCTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACCAAGAGGT	1029
Qy	1324	GATAAACTTTCCACGGTGGGTGGGCCCTCTCAGAAAGAAAAAGAGTCAAGACCTGGCCTA	1383
Db	1030	GATAAACTTTCCACGNNGGTTGGGCCCTCTCAGAAAGAAAAAGAGTCAAGACCTGGCCTN	1089
Qy	1384	GTAACGATPACGATPACCGAGATGACTCCTCTGCAAAAGCCCCAGGGTTAAAGTGCAATCCA	1443
Db	1090	NN	1149
Qy	1444	TCCAAATCTTCCCCGTCAATCCGAGCTTGTCGCCCAAGCTGTGCTGGAGACCTCCCTCTT	1503
Db	1150	NN	1209
Qy	1504	CTTTCAAAATACCTCTACTTAAAAACCTCTCTCTGCAAGGCCCGGTCTCCCATC	1563
Db	1210	NNNNNNNNNNNNCTTACGTTCTCTATTNNAACCTCTCTGCAAGGCCCGGNTCCCATC	1269

QY	1564	AACAGAAAGAGCTCTGTCTCTCGTCTCTCCCAAGCCA	CTTCACTTTTCAAGATGTCG	1623
DB	1270	AACAGAAAGAGGCTNTGTCTCTCGTCTCTCCCAAGCAC	CTTCACTTTTCAAGATNNNN	1329
QY	1624	ATTAGAAACTGGGTGACCCGAAACACCTTCTCATCAC	ACCACCTCACTCCACCTGCTTCG	1683
DB	1330	ATTAGAAACTGGGTGACCCGAAACACCTTCTCATCAC	CCCATCACTCCACCTGCTTCG	1389
QY	1684	GAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCT	CTGAGCCAGAAAGTCATCCCAA	1743
DB	1390	GAGACCAAGATCATGTCTCCAGAAAGCCCTTATTCT	CCTGNNAGCCAGAAAGTCATCCCAA	1449
QY	1744	GCAGAGGCTGCTCTGAGCTAGAAATAGAGTAAGAGG	AGGCTAGACTCAAGCTGCTCG	1803
DB	1450	GCAGAGGCTGCTCTGAGCTAGAAATAGAGTAAGAGG	AGGCTAGACTCAAGCTGCTCG	1509
QY	1804	GAGAGTGTGAAACAAAAGTGTGTGACAGATTGTAACT	TGTGTGACTGAGCTTTGATGGCCAA	1863
DB	1510	GAGAGTGTGAAACAAAAGTGTGTGACAGATTGTAACT	TGTGTGACTGAGCTTTGATGGCCAA	1569
QY	1864	GTTGAAATCTTCAATTTTGGATCTGTGTCGCTTGCTG	GTAAACAGGAGACCTTAGTAAG	1923
DB	1570	GTTGAAATCTTCAATTTTGGATCTGTGTCGCTTGCTG	TGTAAACAGGAGACCTTAGTAAG	1629
QY	1924	GACTCTCTAGTCTCTACCAATCAAGCAAAATTTGAAG	AGAGCTGTAACAGCTATCTCAGAG	1983
DB	1630	GACTCTCTAGTCTCTACCAATCAAGCAAAATTTGAAG	AGAGCTGTAACAGCTATCTCAGAG	1689
QY	1984	CCTCCGCTCTCTATCAGTTCGGTATGCTTCAGAAAGC	TGTGGAAAGCTACCTCTTCTTTTG	2043
DB	1690	CCTCCGCTCTCTGTCAGTTCGGTATGCTTCAGAAAGC	TGTGGAAAGCTACCTCTTCTTTTG	1749
QY	2044	AGACTTTGTGGAGAAAGGCTTGAAATCGTAGGACAA	GAGAAATAGTTTCCCGAGAGATAAA	2103
DB	1750	AGACTTTGTGGAGAAAGGCTTGAAATCGTAGGACAA	GAGAAATAGTTTCCCGAGAGATAAA	1809
QY	2104	AACCTGGTTGTTGGCCATGGCAGCCAAACGGAAGGCT	CAGAAATCATCTCCACGAAAGTCGG	2163
DB	1810	AACCTGGTTGTTGGCCATGGCAGCCAAACGGAAGGCT	CAGAAATCATCTCCACGAAAGTCGG	1869
QY	2164	TCATCCAGACACCCCAATTTCCAGAGACAGAGCGGAA	AGACATTTGCCAGGCCG	2217
DB	1870	TCATCCAGACACCCCAATTTCCAGAGACAGAGCGGAA	AGACATTTGCCAGGCCG	1923

RESULT 10
AL557830

LOCUS	DEFINITION	AL557830	1106 bp	mrna	linear	EST 02-APR-2004
AL557830	Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED	AL557830	1106 bp	mrna	linear	EST 02-APR-2004
	Homo sapiens cDNA clone CS0DJ003YF09 5-PRIME, mrna sequence.					

[illegible]

ACCESSION

ACCESSION
NUMBER

VERSION AL557830.3

VERSION
KEYWORDS
FST

KEYWORDS EST.

NETWORKS	TOP:	HOMO SAPIENS (human)
Source		

SOURCE **Homo sapiens (human)**

ORGANISM
Homo sapiens
Homo sapiens (person)

ORGANISM Homo sapiens

Eukaryota: Metazoa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata

Mammalia: Eutheria: Euarchontoglires: Primates: Catarrhini

Mammalia; Eutheria; Euarchontoglires; Primates; Carnivora

Hominidae: Homo.

Hominidae; Homo.

REFERENCE 1 (bases 1 to 1106)

REFERENCE

1. (bases 1 to 1106)

AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?s=CS0DJ003CC05QP1&c=592.f>.

FEATURES		Location/Qualifiers	
source		1..1106	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CS0DJ003YF09"	
		/cell_type="T CELLS (JURKAT CELL LINE)"	
		/cell_line="JURKAT"	
		/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"	
		10-NORMALIZED"	
		/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
ORIGIN			
Query Match		35.2%; Score 996.2; DB 1; Length 1106;	
Best Local Similarity		99.4%; Pred. No. 1e-265;	
Matches 1031; Conservative		0; Mismatches 3; Indels 3; Gaps 3;	
QY	54	TGCGATTTCTGCTGAACCTGGAGGCAATTTACGACTTTTCTCAGCTGAGGCTTTTCC	113
DB	11	TGCGATTTCTGCTGAACCTGGAGGCAATTTCTCAGCTGAGGCTTTTCC	70
QY	114	TCCGACCTGATGCTCTTCAATTCGGTGTCTCCGCCAGCCACGCTTGGCGTCTTGAGAA	173
DB	71	TCCGACCTGATGCTCTTCAATTCGGTGTCTCCGCCAGCCACGCTTGGCGTCTTGAGAA	130
QY	174	TGGATGCTCTTCAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAA	233
DB	131	TGGATGCTCTTCAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAA	190
QY	234	TGATGAACACACTTCTTATGGAGAAACAGAGTCCCGAGTTCCTCTTTGGATGTACCTT	293
DB	191	TGATGAACACACTTCTTATGGAGAAACAGAGTTCCTCTTTGGATGTACCTT	250
QY	294	CTTTCTGCTCCCAATATGGAACATGTACTAGCAGTTCGCAATGAGAGGCTTTGTCG	353
DB	251	CTTTCTGCTCCCAATATGGAACATGTACTAGCAGTTCGCAATGAGAGGCTTTGTCG	310
QY	354	ATTGTATAACACAGAAATCAAAAGTTTCAAGAAAGAGTGTCTCAAGAAATGGATGCTCA	413
DB	311	ATTGTATAACACAGAAATCAAAAGTTTCAAGAAAGAGTGTCTCAAGAAATGGATGCTCA	370
QY	414	CTGGAATGCGGCTTTTGACCTGGCTGGGTCTCTGGTGAACCTTAAACTTTGTTACAGCAGC	473
DB	371	CTGGAATGCGGCTTTTGACCTGGCTGGGTCTCTGGTGAACCTTAAACTTTGTTACAGCAGC	430
QY	474	AGTGATCAAAACAGCCAAATTTTGGGACGTAAAGCTGTGAGCTGATTTGGAAACATGCAA	533
DB	431	AGTGATCAAAACAGCCAAATTTTGGGACGTAAAGCTGTGAGCTGATTTGGAAACATGCAA	490
QY	534	AGTGATCAATGACGCTCAAGTCAGTGTGCTTTTCTTAAGTTTGAGAAAGCTGTATTCTG	593
DB	491	AGTGATCAATGACGCTCAAGTCAGTGTGCTTTTCTTAAGTTTGAGAAAGCTGTATTCTG	550
QY	594	TACGGGTGGAAGAGATGGCAACATTTATGGTCTGGGATACACAGTGCACAAACAAAGATGG	653
DB	551	TACGGGTGGAAGAGATGGCAACATTTATGGTCTGGGATACACAGTGCACAAACAAAGATGG	610
QY	654	GTTTTATAGGCAAGTGAATCAATTCAGTGGAGCTCAATATACCTCAGACAGCAAAACCCC	713
DB	611	GTTTTATAGGCAAGTGAATCAATTCAGTGGAGCTCAATATACCTCAGACAGCAAAACCCC	670
QY	714	TTCAAAACCCCAAGAGAAACAGAAATTCAGAGGACTTGTCTTCTGTGGATTTCCAGCA	773
DB	671	TTCAAAACCCCAAGAGAAACAGAAATTCAGAGGACTTGTCTTCTGTGGATTTCCAGCA	730
QY	774	AAGTGTTACTGTGCTCTCTTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGA	833
DB	731	AAGTGTTACTGTGCTCTCTTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGA	790

QY	834	TGGGTAATCAAAAGTATGGGATTTAGTGAAGAAATTATCTGCTTATCGACAAGAACCCAT	893
DB	791	TGGGTAATCAAAAGTATGGGATTTAGTGAAGAAATTATCTGCTTATCGACAAGAACCCAT	850
QY	894	AGCATCCAAAGTCTTTCCTGTACCCAGGTAGCAGCACTCGAAAACCTT-GGATATTCAAGTC	952
DB	851	AGCATCCAAAGTCTTTCCTGTACCCAGGTAGCAGCACTCGAAAACCTTGGGATATTCAAGTC	910
QY	953	TGATTTTGATTTCCACTGGCTCTACTTT-ATTGCTAATTGCACAGCACTAACATCTAC	1011
DB	911	TGATTTTGATTTCCACTGGCTCTACTTTTAAATTGCTAATTGCACAGCACTAACATCTAC	970
QY	1012	ATGTTTAAATATGACTGGGTTGAAGACTTCTCAGTGGCTTATTTCATCGACACCAAGAAC	1071
DB	971	ATGTTTAAATATGACTGGGTTGAAGACTTCTCAGTGGCTTATTTCATCGACACCAAGAAC	1029
QY	1072	TCTACCTTTTATGTAAA	1088
DB	1030	TCTACCTTTTATGTAAA	1046
RESULT 11			
AL553783			
LOCUS			
DEFINITION			
AL553783 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
clone CS0DI077YK11 5-PRIME, mRNA sequence.			
ACCESSION			
AL553783			
VERSION			
AL553783.3 GI:45858548			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 998)			
AUTHORS			
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
TITLE			
Full-length cDNA libraries and normalization			
JOURNAL			
Unpublished (2001)			
COMMENT			
On Feb 15, 2001 this sequence version replaced gi:31275597.			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime			
end enriched, double-strand cDNA was digested with Not I and cloned			
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library			
was normalized. Library was constructed by Life Technologies, a			
division of Invitrogen.			
FEATURES		Location/Qualifiers	
source		1..998	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CS0DI077YK11"	
		/tissue_type="Placenta COT 25-NORMALIZED"	
		/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"	
		/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
ORIGIN			
Query Match		35.1%; Score 992.8; DB 1; Length 998;	
Best Local Similarity		99.0%; Pred. No. 8.7e-265;	
Matches 988; Conservative		9; Mismatches 1; Indels 0; Gaps 0;	
QY	1527	TATTAAACCTCTCTGCCAAGCCCGGCTCCCATCAACAGAGAGGCTGTCTCTCTC	1586
DB	1	TATTAAACCTCTCTGCCAAGCCCGGCTCCCATCAACAGAGAGGCTGTCTCTCTC	60
QY	1587	CGTCTCTCCCAAGCACCCTTCATCTTCAAGATGTGATAGAACTGGTGCACCGAAC	1646

Db 61 CGTCTCTCCCAAGCCACCTTCATCTTCAAGATGCGATTAGAAAGTGGTGACCCGAAC 120
QY 1647 ACCTTCTCATCACCAACCCATCACTCCACTGTTGCGGAGACCAAGATCATGTCTCCGAG 1706
Db 121 ACCTTCTCATCACCAACCCATCACTCCACTGTTGCGGAGACCAAGATCATGTCTCCGAG 180
QY 1707 AAAAGCCCTTATTCCTGTGAGCGAGAGTCATCCCAAGCAGAGGCTTGCTCTGAGTCTAG 1766
Db 181 AAAAGCCCTTATTCCTGTGAGCGAGAGTCATCCCAAGCAGAGGCTTGCTCTGAGTCTAG 240
QY 1767 AAATAGAGTAAGAGAGGAGCTAGACTCAAGCTGTCTGAGAGTGTGAACAAAGAGTGTGT 1826
Db 241 AAATAGAGTAAGAGAGGAGCTAGACTCAAGCTGTCTGAGAGTGTGAACAAAGAGTGTGT 300
QY 1827 GAAGAGTTGTAACTGTGACTGAGCTTGATGGCCAAAGTTGAAATCTTTCATTTGGATCT 1886
Db 301 GAAGAGTTGTAACTGTGACTGAGCTTGATGGCCAAAGTTGAAATCTTTCATTTGGATCT 360
QY 1887 GTGCTGCCCTTGCTGAACCGAGAGACCTTAGTAGGACTCTTAGTCTTACCAAAATC 1946
Db 361 GTGCTGCCCTTGCTGAACCGAGAGACCTTAGTAGGACTCTTAGTCTTACCAAAATC 420
QY 1947 AAGCAAAATTTGAAGAGGAGCTGGTACCAGTATCTCAGAGCCTCCGTCTCTTATCAGTCCGTA 2006
Db 421 AAGCAAAATTTGAAGAGGAGCTGGTACCAGTATCTCAGAGCCTCCGTCTCTTATCAGTCCGTA 480
QY 2007 TGCCTCAGAAAGCTGTGGAACGCTACCTCTTCTTGTGAGACCTTGTGGAGAAAGGCTCTGA 2066
Db 481 TGCCTCAGAAAGCTGTGGAACGCTACCTCTTCTTGTGAGACCTTGTGGAGAAAGGCTCTGA 540
QY 2067 AATGTTAGGCAAGAGAAATAGTTTCCCGAGAGATTAAGGACTCTTAGTCTTACCAAAATC 2126
Db 541 AATGTTAGGCAAGAGAAATAGTTTCCCGAGAGATTAAGGACTCTTAGTCTTACCAAAATC 600
QY 2127 CAAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCCGTATCCAGACACCCAAATCCAG 2186
Db 601 CAAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCCGTATCCAGACACCCAAATCCAG 660
QY 2187 GAGACAGCGGAAAGACATTCGCAAGCCCGTCAACATCAGCCGAGCTCATGAGGAA 2246
Db 661 GAGACAGCGGAAAGACATTCGCAAGCCCGTCAACATCAGCCGAGCTCATGAGGAA 720
QY 2247 AATCTGCACATCTTCATAGAAAGTCCAGAGGAGCTTCTGTGCTCGACACTCAAC 2306
Db 721 AATCTGCACATCTTCATAGAAAGTCCAGAGGAGCTTCTGTGCTCGACACTCAAC 780
QY 2307 AGAATATAGATTTCTAATCTGAGTGAATTAAGTCTGAGCTTTGGTCCACTAAAACAGCTGAG 2366
Db 781 AGAATATAGATTTCTAATCTGAGTGAATTAAGTCTGAGCTTTGGTCCACTAAAACAGCTGAG 840
QY 2367 CTTTGGTCCACTAAAACAGAGTGAATAATACAAAGAGTGAATTAATCTTGGTCTTTAA 2426
Db 841 CTTTGGTCCACTAAAACAGAGTGAATAATACAAAGAGTGAATTAATCTTGGTCTTTAA 900
QY 2427 GAAGCTGCTTTCTAATTTTGAACAAATCTTTTCAACGCTGAAATGTACCTAATCTGG 2486
Db 901 GAAGCTGCTTTCTAATTTTGAACAAATCTTTTCAACGCTGAAATGTACCTAATCTGG 960
QY 2487 TTCTACTACATAATATATATGAGCTTCCCGAGGATG 2524
Db 961 TTCTACTACATAATATATGAGCTTCCCGAGGATG 998

RESULT 12
AL576790/c 1042 bp mRNA linear EST 07-APR-2004
LOCUS
DEFINITION AL576790 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D107YK11 3-PRIME, mRNA sequence.
ACCESSION AL576790
VERSION AL576790.3 GI:46255889
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchntoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1042)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:31315071.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1..1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D107YK11"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone.Lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

ORIGIN

Query Match 34.5%; Score 976.4; DB 1; Length 1042;
Best Local Similarity 97.7%; Pred. No. 3.3e-260; Indels 2; Gaps 2;
Matches 1000; Conservative 9; Mismatches 13;
QY 1767 AAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCTGAGAGTGTGAACAAAGTGTGT 1826
Db 1022 AGAATRGRTAAAGRGCGCTAACTCAAGCTTCTGGAGAGTG-KAAACAAAGTGTGT 964
QY 1827 GAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAATCTTTCATTTGGATCT 1886
Db 963 GAAGAG-TGTAACCTGTGTGACTGAGCTTGATGSCCAAGTTGAAATCTTTCATTTGGATCT 905
QY 1887 GTGCTGCCCTTCTGTTAACCCAGGAGACCTTAGTAGAGGACTCTTAGTCTTACCAAAATC 1946
Db 904 GTGCTGCCCTTCTGTTAACCCAGGAGACCTTAGTAGAGGACTCTTAGTCTTACCAAAATC 845
QY 1947 AAGCAAAATTTGAAGAGGCTGGTACCAGTATCTCAGAGGCTCCGTCTCTATCAGTCCGTA 2006
Db 844 AAGCAAAATTTGAAGAGGCTGGTACCAGTATCTCAGAGGCTCCGTCTCTATCAGTCCGTA 785
QY 2007 TGTCTCAGAAAGCTGTGAAACGCTACCTCTTCTTTTGTGAGACCTTGTGAGAAAGGCTCTGA 2066
Db 784 TGTCTCAGAAAGCTGTGAAACGCTACCTCTTCTTTTGTGAGACCTTGTGAGAAAGGCTCTGA 725
QY 2067 AATGTTAGGCAAGAGAAATAGTTTCCCGAGAGATTAAGGACTCTTAGTCTTACCAAAATC 2126
Db 724 AATGTTAGGCAAGAGAAATAGTTTCCCGAGAGATTAAGGACTCTTAGTCTTACCAAAATC 665
QY 2127 CAAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCCGTATCCAGACACCCAAATCCAG 2186
Db 664 CAAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCCGTATCCAGACACCCAAATCCAG 605
QY 2187 GAGACAGCGGAAAGACATTCGCAAGCCCGGTACCAATCAGCCCGAGTCCATCAGGAA 2246
Db 604 GAGACAGCGGAAAGACATTCGCAAGCCCGGTACCAATCAGCCCGAGTCCATCAGGAA 545
QY 2247 AATCTGCACATCTTCCATAGAAAGTCCAGAGGAGCTTCTGTGCTCTGACACTCAAC 2306
Db 544 AATCTGCACATCTTCCATAGAAAGTCCAGAGGAGCTTCTGTGCTCTGAACTCAAC 485
QY 2307 AGAATATAGATTTCTAATCTGAGTGAATTAAGTCTGAGCTTTGGTCCACTAAAACAGTCTGAG 2366

CONSRMT
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

RIKEN Genome Exploration Research Group
Antisense transcription in the mammalian transcriptome
Science 309 (5740), 1564-1566 (2005)
16141073

8 (bases 1 to 4308)
Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K.,
Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S.,
Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N.,
Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D.,
Shibata, K., Shiraki, T., Tagami, W., Tagami, Y., Waki, K., Watahiki, A.,
Muramatsu, M., and Hayashizaki, Y.

Direct Submission
Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa, 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES
source
Location/Qualifiers
1..4308
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:E130114J22"
/db_xref="taxon:10090"
/clone="E130114J22"
/tissue_type="eyeball"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
1..4308
misc_feature
/note="L2DTL protein homolog [Mus musculus]
(UniProt|Q8BM38, evidence: FASTA, 91.9%ID, 52.1%length,
match=1145)
putative"

ORIGIN

Query Match 31.4%; Score 890; DB 6; Length 4308;
Best Local Similarity 87.1%; Pred. No. 6.1e-236;
Matches 1002; Conservative 0; Mismatches 140; Indels 8; Gaps 2;

QY 12 GAGTCGAGCGATACGATTGTTGTGTGAGAGCGCAACGTGCGATTCTCTGCTGAAT 71
DB 8 GAGTGTGAGCGAAACCGCGCGGTGTGAGGGTGCAGACGGCGCTCTGTGGAAT 67
QY 72 TGAGGCG-ATTTCTAGACATTTTCTCTCAGCTGAGGCTTTTCTCGACCGCTGATGCTCT 130
DB 68 TGAGGCGGATCTTCTCGGCGTTCCTCTCAGTGCCTTCTCTCAGGCGCAGATGCTCT 127
QY 131 TCAATTGCGGTGCTCGCGACGCCAGCTTGGCGTCTCTGAGAAATGAGATGCTTCTCAAT 190
DB 128 TCAACTCGGTACTCTCGCGACGCCAGCTCGCGCTCTCTGAGGAAACGGGTGCTTCAAT 187
QY 191 ACCCTCTCAATCCCTCTGACTGTTATTCAGTGCAGTGTGATGATGACACATCTCT 250
DB 188 ACCCTCTCAATCCCTCTTAAGTGGTATTCAGTGCACACTGTAAAGATGAACACAGCTCT 247
QY 251 ATGAGAAACAGGAGTCCCAAGTTCCTCTTTGGATGTACCTTCTCTCTGCTCCCAATA 310
DB 248 ATGAGAAACAGAGTCCCAAGTTCCTCTTTGGATGCACCTCTGTACCGCTCCAGTA 307
QY 311 TGAACATGATCTAGCAGTTGCCAATGAAGAAGCTTTGTTGCATGTTGATTAACACAGAT 370
DB 308 TGGAGCATATATTAGCAGTTGCTAAATGAAGAAGCTTTGTCAGATATTATAATACAGAT 367

QY 371 CACAAAGTTTTCAGAAAGAGTGTCTTCAAGAATGGATGGCTCACTGGAATGCCGTCTTTG 430
DB 368 CACAAACTAGCAAAAGAGCATGCTTCAAGGAGTGGATGGCTCACTGGAATGCTGCTTTG 427
QY 431 ACTGGCTGGGTTCCTGGTGAACTTAAACTTTTACAGCAGCAGGTGATCAAAACAGCCA 490
DB 428 ACTTGGCTGGGTTCCTGGTGAACTTAAACTT-----CAGCCGGTGATCAGACAGCCA 480
QY 491 AATTTTGGGACCTTAAAGCTGTGAGCTGTGAGTGGACATGCAAGAGTCAATCAATGCAGCC 550
DB 481 AATTTTGGGATGTAAAGCTGTGAGCTGTGAGTGGGACATGCAAGAGCCACCACTGCGCC 540
QY 551 TCAAGTCAGTTCCTTTTCTAAGTTTGAAGAGCTGTATTTCTGTACGGGTGGAAGAGATG 610
DB 541 TCAAGTCTGTAGCTTTTCCCAAGTTTCAAAAAGCTGTGTTCTCTACAGGGGGAGAGAGC 600
QY 611 GCAACATTTATGTCCTGGGATACAGGTGCAACAAAAAGATGGTTTTATAGCAAGTGA 670
DB 601 GCAACATTTATGATCTGGGACACAGGTGTAAACAAAAAGATGGTTTTATAGCAAGTGA 660
QY 671 ATCAATCAGTCGAGCTCACAATACCTCAGACAGCAACCAACCCCTTCABAACCCCAAGA 730
DB 661 ATCAATCAGTCGAGCTCACAATACCTCAGACAGCAACCAACCCCTTCABAACCCCAAGA 720
QY 731 AACAGAAATTCAAAAGGACTTGTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGCTCC 790
DB 721 AACAAATTCAAAAGGACTTGTCTCTCTGTGGATTTCCAGCAGAGTGTACTGTGCTCC 780
QY 791 TCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATTAATCAAAAGTAT 850
DB 781 TCTTTCAAGATGAGAAATACATTTAGTCTCAGCAGGAGCGGTGGATGGAATTAATCAAAAGTAT 840
QY 851 GGGATTTTACGTAAGAAATATACCTGCTTTCGACAGAGAACCCATAGCATCCAAAGTCTTTCC 910
DB 841 GGGATTTTGGCAGAAATATACCTGCTTATCGACAGAGAACCCATAGCATCCAAAGTCTTTCC 900
QY 911 TGTAACCCAGTACGAGCACTCGAAACCTTGGATATTTCAAGTCTGATTTTGGATTTCCACTG 970
DB 901 TGTAACCCAGTACGAGCACTCGAAAGCTAGGATATCTCGAGTTTGGTTTAGACTCTACTG 960
QY 971 GCTCTACTTTATTTGCTAATTCGACAGAGTAAATCTACATCTATATGTTTATATGACTGGGT 1030
DB 961 GCTCTACTTTATTTGCTAATTCGACAGATGACAAATCTATATGTTTCAATATGACTGGGT 1020
QY 1031 TGAAGACTTCTCCAGTGGCTATTTTCAATGGACCCAGCACTCTACCTTTTATGTTAAAT 1090
DB 1021 TAAAGACTTCTCCGCTGGCTGTCTTCAATGGACCCAGCACTCTACCTTTTATGTTAAAT 1080
QY 1091 CCAGCTTATGTCAGATGACCACTTTTATGTCAGTGGCTCAAGTGTGATGAAGCTGCCTACA 1150
DB 1081 CAAGCTTATGTCAGATGACCACTTTTATGTCAGTGGTTCAGTGTGATGAAGCTGCCTACA 1140
QY 1151 TATGGAAGGT 1160
DB 1141 TTTGGAAGGT 1150

RESULT 15
BQ431493
LOCUS
DEFINITION BQ431493 949 bp mRNA linear EST 24-MAY-2002
AGENCOURT 7831396 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6152134
5', mRNA sequence.
ACCESSION BQ431493
VERSION BQ431493.1 GI:21170579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 949)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13489 row: h column: 23

High quality sequence stop: 639.

FEATURES
source

Location/Qualifiers

1..949

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6152134"

/tissue_type="retinoblastoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.75 kb. Library constructed by Life

Technologies."

ORIGIN

Query Match 31.3%; Score 885.8; DB 3; Length 949;
Best Local Similarity 98.2%; Pred. No. 6e-235;
Matches 927; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

QY 1465 GCAGCTTGCCCAAGCTGTGTGAGACCTCCCTCTTCTTCAAAATCTCTAGGTTTC 1524
DB 1 GCAGCTTGCCCAAGCTGTGTGAGACCTCCCTCTTCTTCAAAATCTCTAGGTTTC 60

QY 1525 TCTATTAAACCTCTCTCCAGGCGGTCTCCATCAAGAGAGGCTCTGTCTCC 1584
DB 61 TCTATTAAACCTCTCTCCAGGCGGTCTCCATCAAGAGAGGCTCTGTCTCC 120

QY 1585 TCCGTCTCTCCCAAGCAGCTTCTCATCTTCAAGATGTCGATTAGAACTGGGTGACCCGA 1644
DB 121 TCCGTCTCTCCCAAGCAGCTTCTCATCTTCAAGATGTCGATTAGAACTGGGTGACCCGA 180

QY 1645 ACACCTTCTCTATCACCAACCATCATCTCACTGCTTGGAGACCAAGATCATGTCTCG 1704
DB 181 ACACCTTCTCTATCACCAACCATCATCTCACTGCTTGGAGACCAAGATCATGTCTCG 240

QY 1705 AGAAAGCCCTTATCTCTGTGAGCAGAGTCTATCCCAAGCAGAGGCTTGTCTGAGTCT 1764
DB 241 AGAAAGCCCTTATCTCTGTGAGCAGAGTCTATCCCAAGCAGAGGCTTGTCTGAGTCT 300

QY 1765 AGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAGTGT 1824
DB 301 AGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAGTGT 360

QY 1825 GTGAAGAGTTGTAACCTGTGAGTCTGAGCTTGTGAGGCAAGTGTGAAATCTTCAATTGGAT 1884
DB 361 GTGAAGAGTTGTAACCTGTGAGTCTGAGCTTGTGAGGCAAGTGTGAAATCTTCAATTGGAT 420

QY 1885 CTGTGCTGCTTGTCTGTAACCAAGGAGACCTTTAGTAAGGACTCTCTAGGTCTCTACCAA 1944
DB 421 CTGTGCTGCTTGTCTGTAACCAAGGAGACCTTTAGTAAGGACTCTCTAGGTCTCTACCAA 480

QY 1945 TCAAGCAAAATTGAAGAGAGTGGTACAGTATCTCAGAGCCTCCGCTCTCTATCAGTCCG 2004
DB 481 TCAAGCAAAATTGAAGAGAGTGGTACAGTATCTCAGAGCCTCCGCTCTCTATCAGTCCG 540

QY 2005 TATGCTTTCAGAAAGCTGTGGAACGCTACCTCTTCTTTGAGACCTTGTGGAAGGGTCT 2064
DB 541 TATGCTTTCAGAAAGCTGTGGAACGCTACCTCTTCTTTGAGACCTTGTGGAAGGGTCT 600

QY 2065 GAAATGCTAGGCAAGAGAGTACTCTCCAGAGAAATAAAACCTGGTGTGTCGCATGGCA 2124

Search completed: November 9, 2006, 09:59:09

Job time : 13056 secs

DB 601 GAAATGGTAGGCAAGAGAGATAGTTCCCGAGAGAAATAAAACCTGGTGTGTCGCATGGCA 660
QY 2125 GCCAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTATCCAGACACCCAAATTC 2184
DB 661 GCCAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTATCCAGACACCCAAATTC 720
QY 2185 AGGAGACAGAGCGGGAAGAGACATTCGCAAGCCCGGTCAACATCAAGCCAGCTCCATGAGG 2244
DB 721 AGGAGACAGAGCGGGAAGAGACATTCGCAAGCCCGGTCAACATCAAGCCAGCTTCATGANG 780
QY 2245 AAAATCTGCACATCTTCCATAGAAAGT-CCAGAGGAGTCTTCTGTGCTCTGAACACTC 2303
DB 781 AAAATCTGCACATCTTCCATAGAAAGTCCCGAGGAGACTTCTGTGCTCTGAACACTC 840
QY 2304 AACAGAAATTAAGATTTCTAATCTGAGTGAGTTACT-GAGCTTTGCTCCACTAAAAACAAGC 2362
DB 841 AACAGAAATTAAGATTTCTAATCTGAGTGAGTTACTCTGGAGCTTTGGTCCACTAAAAACAAGC 900

QY 2363 TGA--GCTTTGGTCCACTTAAACAAAGATGAAATAATACAAGAGTG 2404
DB 901 TGAAGCTTTGGTCCACTTAAACAAAGATGAAATAATACCAGAGTG 944

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 06:20:51 ; Search time 643 Seconds

(without alignments)
8714.811 Million cell updates/sec

Title: US-10-726-160-1

Perfect score: 2831

Sequence: 1 ggcacagcggagtggtggag.....atgaaaaaaaaaaaaaa 2831

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2471159 seqs, 989689746 residues

Total number of hits satisfying chosen parameters: 4942318

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New.*

- 1: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 2: /EMC Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /EMC Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /EMC Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 5: /EMC Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 6: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 8: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 9: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 10: /EMC Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	2831	100.0	2831	8	US-11-266-748A-23628 Sequence 23628, A
2	2789.8	98.5	4221	8	US-11-266-748A-23288 Sequence 23288, A
3	2789.8	98.5	4221	8	US-11-266-748A-56960 Sequence 56960, A
4	2442.4	86.3	4083	6	US-10-540-898-673 Sequence 673, App
5	1937.4	68.4	1949	8	US-11-266-748A-184890 Sequence 184890
6	1302.2	46.0	2436	6	US-10-540-898-670 Sequence 670, App
7	1235	43.6	1270	8	US-11-266-748A-184891 Sequence 184891
8	1235	43.6	1270	8	US-11-266-748A-192385 Sequence 192385
9	835.8	29.5	8892	6	US-10-540-898-672 Sequence 672, App
10	802	28.3	931	8	US-11-266-748A-93508 Sequence 93508, A
11	802	28.3	931	8	US-11-266-748A-146319 Sequence 146319
12	799.4	28.2	883	8	US-11-266-748A-41156 Sequence 41156, A
13	799.4	28.2	883	8	US-11-266-748A-210423 Sequence 210423
14	700.8	24.8	798	8	US-11-266-748A-261935 Sequence 261935
15	700.8	24.8	798	8	US-11-266-748A-322452 Sequence 322452
16	649.8	23.0	57860	6	US-10-540-898-669 Sequence 669, App
17	613	21.7	745	6	US-10-540-898-675 Sequence 675, App
18	495.2	17.5	500	8	US-11-266-748A-9970 Sequence 9970, App
19	439.4	15.5	485	8	US-11-266-748A-262502 Sequence 262502
20	439.4	15.5	485	8	US-11-266-748A-323019 Sequence 323019
21	291.4	10.3	464	6	US-10-513-369-1049 Sequence 1049, App
22	275.8	9.7	379	8	US-11-266-748A-232873 Sequence 232873
23	103.6	3.7	1914	9	US-11-218-305-21058 Sequence 21058, A

ALIGNMENTS

RESULT 1

US-11-266-748A-23628

; Sequence 23628, Application US/11266748A

; Publication No. US20060134663A1

; GENERAL INFORMATION:

; APPLICANT: Harkin, Paul

; APPLICANT: Johnston, Patrick

; APPLICANT: Mulligan, Karl

; TITLE OF INVENTION: Transcriptome Microarray Technology and

; FILE REFERENCE: 55815-0102 (319189)

; CURRENT APPLICATION NUMBER: US/11/266, 748A

; PRIOR FILING DATE: 2005-11-03

; PRIOR APPLICATION NUMBER: EP 04105479.2

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105482.6

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105483.4

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105507.0

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105485.9

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105484.2

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: US 60/662,276

; PRIOR FILING DATE: 2005-03-14

; PRIOR APPLICATION NUMBER: US 60/700,293

; PRIOR FILING DATE: 2005-07-18

; NUMBER OF SEQ ID NOS: 483996

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 23628

; LENGTH: 2831

; TYPE: DNA

; ORGANISM: Homo Sapiens

US-11-266-748A-23628

Query Match 100.0%; Score 2831; DB 8; Length 2831;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2831; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACAGCGGAGTTGGAGCGGATACGATTTGTTGTGAGCGGCAACGTGGCAT 60

DB 1 GGCACAGCGGAGTTGGAGCGGATACGATTTGTTGTGAGCGGCAACGTGGCAT 60

QY 61 TCTGCTGACTTTGGAGGCATTTCTACGACTTTTCTCTCAGTCAGGCTTTTCTCCGACC 120

Sequence 13367, A
Sequence 75767, A
Sequence 258, App
Sequence 532, App
Sequence 144, App
Sequence 272, App
Sequence 202830, A
Sequence 178827, A
Sequence 85525, A
Sequence 138336, A
Sequence 228469, A
Sequence 25251, A
Sequence 257849, A
Sequence 318366, A
Sequence 355179, A
Sequence 385180, A
Sequence 438558, A
Sequence 21862, A
Sequence 31052, A
Sequence 78160, A
Sequence 75238, A
Sequence 108274, A

61	Db	 TCTGCTGAACATTGGAGGCAATTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCTCGAGCC	120
121	Qy	 CTGATGCTCTTCAATTCGGTGCTCCGGCAGGCCAGCTTGGCGCTCTGAGAAATGGATGG	180
121	Db	 CTGATGCTCTTCAATTCGGTGCTCCGGCAGGCCAGCTTGGCGCTCTGAGAAATGGATGG	180
181	Qy	 TCTTCACAATAACCCCTCTTCAATCCCTCTTGACTGGTTATCAGTGCAGTGGTAATGATGAA	240
181	Db	 TCTTCACAATAACCCCTCTTCAATCCCTCTTGACTGGTTATCAGTGCAGTGGTAATGATGAA	240
241	Qy	 CACACTTCTTATGGAGAAACAGGAGTCCAGTTCCTCTTTTGGATGTACTCTCTCTCT	300
241	Db	 CACACTTCTTATGGAGAAACAGGAGTCCAGTTCCTCTTTTGGATGTACTCTCTCTCTCT	300
301	Qy	 GCTCCCAATATGGACATGTACTAGCAGTTGCCAATGAGNAGGCTTTGTTTCGATGTAT	360
301	Db	 GCTCCCAATATGGACATGTACTAGCAGTTGCCAATGAGNAGGCTTTGTTTCGATGTAT	360
361	Qy	 AACACAGAAATCAAAAGTTTCAGAAAAAGGTCTTCAAGAATGGATGGCTCACTGGAAAT	420
361	Db	 AACACAGAAATCAAAAGTTTCAGAAAAAGGTCTTCAAGNAATGGATGGCTCACTGGAAAT	420
421	Qy	 GCCGTCTTTTGACCTGGGCTCGGTTCTCTGGTGAACCTTAAACTTGTACAGCAGCAGTGTAT	480
421	Db	 GCCGTCTTTTGACCTGGGCTCGGTTCTCTGGTGAACCTTAAACTTGTACAGCAGCAGTGTAT	480
481	Qy	 CAAAACAGCCAAATTTTGGGACGTAAAAGCTGGTGAGCTGATTTGAAACATGCAAAAGTCTAT	540
481	Db	 CAAAACAGCCAAATTTTGGGACGTAAAAGCTGGTGAGCTGATTTGAAACATGCAAAAGTCTAT	540
541	Qy	 CAATGACGCTCAAGTCAGTTGCTTTTCTAAGTTTCAGAAAGCTGTATTTCTGTACGGGT	600
541	Db	 CAATGACGCTCAAGTCAGTTGCTTTTCTAAGTTTCAGAAAGCTGTATTTCTGTACGGGT	600
601	Qy	 GGAAAGAGATGGCAACATTTATGGTCTGGGATACCAAGTGCACAAAAAAGATGGGTTTTAT	660
601	Db	 GGAAAGAGATGGCAACATTTATGGTCTGGGATACCAAGTGCACAAAAAAGATGGGTTTTAT	660
661	Qy	 AGGCAAGTAATCAAAATCAGTGGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCAAAA	720
661	Db	 AGGCAAGTAATCAAAATCAGTGGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCAAAA	720
721	Qy	 CCCAAGAGAAACAGAAATTCAAAAGGACTTCTCTCTCTGTGGATTTCCACGACAAAGTGT	780
721	Db	 CCCAAGAGAAACAGAAATTCAAAAGGACTTCTCTCTCTGTGGATTTCCACGACAAAGTGT	780
781	Qy	 ACTGTGGTCTCTTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATA	840
781	Db	 ACTGTGGTCTCTTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATA	840
841	Qy	 ATCAAAAGTATGGGATTTTACGTAAAGAAATTAATCTGCTTATGCAAGAACCCATAGCATCC	900
841	Db	 ATCAAAAGTATGGGATTTTACGTAAAGAAATTAATCTGCTTATGCAAGAACCCATAGCATCC	900
901	Qy	 AAGTCTTTCTGTATCCAGGATAGCAGCATCTGCAAACTTCGGATTTCAAGCTCGATTTTCG	960
901	Db	 AAGTCTTTCTGTATCCAGGATAGCAGCATCTGCAAACTTCGGATTTCAAGCTCGATTTTCG	960
961	Qy	 GATTCCACTGGCTCTACTTTTATTTGCTAAATGTCACAGACGATACATCTACATGTTTAAT	1020
961	Db	 GATTCCACTGGCTCTACTTTTATTTGCTAAATGTCACAGACGATACATCTACATGTTTAAT	1020
1021	Qy	 ATGACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGAACACAGAACTCTACCTTTT	1080
1021	Db	 ATGACTGGGTTGAAGACTTCTCTCAGTGGCTATTTTCAATGAACACAGAACTCTACCTTTT	1080
1081	Qy	 TATGTAAATCCAGCCTTTAGTCCAGATGACCAAGTTTTTATGTACGTGGCTCAAGTGATGAA	1140
1081	Db	 TATGTAAATCCAGCCTTTAGTCCAGATGACCAAGTTTTTATGTACGTGGCTCAAGTGATGAA	1140
1141	Qy	 GCTGCTACTATATGGAAGGCTCCACACCTCGGCAACCTCCTACTGTGCTCTGGGTCTAT	1200

Db	1141	GCTGCCCTACATATGGAAGGCTCTCCA	CACCCTGGGAACCTCTCTACTGTGCTCTCTGGGTCAT	1200
Qy	1201	TCTCAAGAGGTCACGCTCTGTGTGCT	3GTGTGCCATCTGACTTTCAAAAGATTGCTACCTGT	1260
Db	1201	TCTCAAGAGGTCACGCTCTGTGTGCT	GGTGTCCATCTGACTTTACAAGATTGCTACCTGT	1260
Qy	1261	TCTGATGACAATACACTAAAAATCT	TGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGA	1320
Db	1261	TCTGATGACAATACACTAAAAATCT	TGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGA	1320
Qy	1321	GGTGATAAATCTTTCCACGGTGGGT	GGGCGCTCTCAGAGAAAGAGTCAAGACCTGGC	1380
Db	1321	GGTGATAAATCTTTCCACGGTGGGT	GGGCGCTCTCAGAGAAAGAGTCAAGACCTGGC	1380
Qy	1381	CTAGTAAACAGTAACGAGTAGCAGAGT	ACTCTGCCAAAGCCCCCAGGGTAAAGTGCAT	1440
Db	1381	CTAGTAAACAGTAACGAGTAGCAGAGT	ACTCTGCCAAAGCCCCCAGGGTAAAGTGCAT	1440
Qy	1441	CCATCCAAATCTTTCCCGGTCAAT	CCGCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCT	1500
Db	1441	CCATCCAAATCTTTCCCGGTCAAT	CCGCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCT	1500
Qy	1501	CTTCCCTTCAAATACTCTCTA	GGTTCCTATTTAAACCTTCTCTGCCAAGGCCCGGTCTCCC	1560
Db	1501	CTTCCCTTCAAATACTCTCTA	GGTTCCTATTTAAACCTTCTCTGCCAAGGCCCGGTCTCCC	1560
Qy	1561	ATCAACAGAGAGGCTCTGTCTCT	CCGCTCTCTCCCAAGCCACTTCACTTTCAAGATG	1620
Db	1561	ATCAACAGAGAGGCTCTGTCTCT	CCGCTCTCTCCCAAGCCACTTCACTTTCAAGATG	1620
Qy	1621	TCGATTAGAAACTGGGTGAC	CCGAACCTTCTCATCACCCCATCACTCCACCTGT	1680
Db	1621	TCGATTAGAAACTGGGTGAC	CCGAACCTTCTCATCACCCCATCACTCCACCTGT	1680
Qy	1681	TCGGAGACCAAGATCATGTCT	CCGAGAAAAGCCCTTATTCCTGTGAGCCAGAGTCAATCC	1740
Db	1681	TCGGAGACCAAGATCATGTCT	CCGAGAAAAGCCCTTATTCCTGTGAGCCAGAGTCAATCC	1740
Qy	1741	CAAGCAGAGGCTTGCTCTGAGTCT	AGAAATAGAGTAAAGAGGAGGCTTAGACTCAAGCTGT	1800
Db	1741	CAAGCAGAGGCTTGCTCTGAGTCT	AGAAATAGAGTAAAGAGGAGGCTTAGACTCAAGCTGT	1800
Qy	1801	CTGGAGAGTGTGAACAAAAAGTGT	GGAAGTTGTAACTGTGTGACTGAGCTTGATGGC	1860
Db	1801	CTGGAGAGTGTGAACAAAAAGTGT	GGAAGTTGTAACTGTGTGACTGAGCTTGATGGC	1860
Qy	1861	CAAGTTGAAAATCTTCAATTTGGAT	CTGTGCTGCCCTGCTGTTAACCGGAAGACCTTAGT	1920
Db	1861	CAAGTTGAAAATCTTCAATTTGGAT	CTGTGCTGCCCTGCTGTTAACCGGAAGACCTTAGT	1920
Qy	1921	AAGGACTCTCTAGTCTCTACCAAT	CAAGCAAAATGAAGGAGCTGTGTACAGTATCTCA	1980
Db	1921	AAGGACTCTCTAGTCTCTACCAAT	CAAGCAAAATGAAGGAGCTGTGTACAGTATCTCA	1980
Qy	1981	GAGCCTCCGTCTCTATCAGTCCGT	ATGCTTTCAGAAAGCTGTGGAAGCTTACCTTTCCCT	2040
Db	1981	GAGCCTCCGTCTCTATCAGTCCGT	ATGCTTTCAGAAAGCTGTGGAAGCTTACCTTTCCCT	2040
Qy	2041	TTGAGACCTTTGTGAGAGGGTCT	GAAATGGTATGGCAAGAGATAGTTTCCCGAGAGAT	2100
Db	2041	TTGAGACCTTTGTGAGAGGGTCT	GAAATGGTATGGCAAGAGATAGTTTCCCGAGAGAT	2100
Qy	2101	AAAAACTGGTTGTGGCCATGGC	AGCCAAACGGAAGCTGAGATCCATCTCCACGAAGT	2160
Db	2101	AAAAACTGGTTGTGGCCATGGC	AGCCAAACGGAAGCTGAGATCCATCTCCACGAAGT	2160
Qy	2161	CCGTCAATCCAGACACCCCAAT	TCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTC	2220
Db	2161	CCGTCAATCCAGACACCCCAAT	TCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTC	2220
Qy	2221	ACCATCAAGCCAGCTCCATGAG	AAATCTGCACTACTTCNTAGAAAGTCCAGGAG	2280
Db	2221	ACCATCAAGCCAGCTCCATGAG	AAATCTGCACTACTTCNTAGAAAGTCCAGGAG	2280

```
QY 2281 GACTTCTGTGGTCCCTGAACACCTCAACAGAAATTATAGATTTCTAATCTGAGTGAAGTTACTGA 2340
Db |||||||
QY 2281 GACTTCTGTGGTCCCTGAACACCTCAACAGAAATTATAGATTTCTAATCTGAGTGAAGTTACTGA 2340
Db |||||||
QY 2341 GCTTTGGTCCACTAAACAAAGCTGAGCTTTGGTCCACTAAACAAAGATGAAAAATACAAAG 2400
Db |||||||
QY 2341 GCTTTGGTCCACTAAACAAAGCTGAGCTTTGGTCCACTAAACAAAGATGAAAAATACAAAG 2400
Db |||||||
QY 2401 AGTGACTCTATACCTGCTCTTTAAGAAAGCTGCTTTTCAATTTTATAGACAAAATCTTT 2460
Db |||||||
QY 2401 AGTGACTCTATACCTGCTCTTTAAGAAAGCTGCTTTTCAATTTTATAGACAAAATCTTT 2460
Db |||||||
QY 2461 TCAACGCTGAAATGTACCTAACTGTGTTCTACTACCATAATGTATATGAGCTTCCCGAG 2520
Db |||||||
QY 2461 TCAACGCTGAAATGTACCTAACTGTGTTCTACTACCATAATGTATATGAGCTTCCCGAG 2520
Db |||||||
QY 2521 GATGAATGCTGTGTTTAAATTTTATATAAGTAAATTTGTCACTTAGCAATTTTGAATGAAT 2580
Db |||||||
QY 2521 GATGAATGCTGTGTTTAAATTTTATATAAGTAAATTTGTCACTTAGCAATTTTGAATGAAT 2580
Db |||||||
QY 2581 AGCTTTCACCTTTTAAATTTTATATAAGTAAATTTGTCACTTAGCAATTTTGAATGAAT 2640
Db |||||||
QY 2581 AGCTTTCACCTTTTAAATTTTATATAAGTAAATTTGTCACTTAGCAATTTTGAATGAAT 2640
Db |||||||
QY 2641 CAAAAACAAAGTTTCTTTGTTATCTGAAACTTTCTATGCTCAGTGGAAGTATCTGCCAG 2700
Db |||||||
QY 2641 CAAAAACAAAGTTTCTTTGTTATCTGAAACTTTCTATGCTCAGTGGAAGTATCTGCCAG 2700
Db |||||||
QY 2701 CCACAGCATGAGGCTGTGAGGCTGACGTGAGAAATCTCTCTGCTGAAGACCCCTGGTTCT 2760
Db |||||||
QY 2701 CCACAGCATGAGGCTGTGAGGCTGACGTGAGAAATCTCTCTGCTGAAGACCCCTGGTTCT 2760
Db |||||||
QY 2761 GTTCTGCTCCACATGATATATTTTATTTGAAATACATAATCTTTTCACTATGAAAAA 2820
Db |||||||
QY 2761 GTTCTGCTCCACATGATATATTTTATTTTGAATACATAATCTTTTCACTATGAAAAA 2820
Db |||||||
QY 2821 AAAAAA 2831
Db |||||||
QY 2821 AAAAAA 2831
Db |||||||
```

RESULT 2

```
US-11-266-748A-23288
; Sequence 23288, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
```

```
; SEQ ID NO 23288
; LENGTH: 4221
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-23288
```

```
Query Match          98.5%; Score 2789.8; DB 8; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 22 CGATACCAATTTGTTGTTGAGAGCGCAAGCTGGGATTTCTGCTGAACACTTTGGAGGCATTT 81
Db |||||||
QY 82 TCTAGCACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 141
Db |||||||
QY 61 TCTAGCACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 120
Db |||||||
QY 142 CTCCGCCAGCCCGAGCTTTGGGCTCTCTGAGAAATGGATGGTCTTCAAAATACCTCTTCAA 201
Db |||||||
QY 121 CTCCGCCAGCCCGAGCTTTGGGCTCTCTGAGAAATGGATGGTCTTCAAAATACCTCTTCAA 180
Db |||||||
QY 202 TCCCTTCTGACTGGTTATCAGTGCAGTGTAAATGATGAACACACTTCTTATGGAGAAACA 261
Db |||||||
QY 181 TCCCTTCTGACTGGTTATCAGTGCAGTGTAAATGATGAACACACTTCTTATGGAGAAACA 240
Db |||||||
QY 262 GGAGTCCAGCTTCTCTCTTTGGGATGTACCTTCTCTGCTCCCAATATGGAACATGTA 321
Db |||||||
QY 241 GGAGTCCAGCTTCTCTCTTTGGGATGTACCTTCTCTGCTCCCAATATGGAACATGTA 300
Db |||||||
QY 322 CTAGCAGTTGCAATGAAAGAGGCTTTGTTGATTTGTTATTAACACAGAAATCAAAAGTTTC 381
Db |||||||
QY 301 CTAGCAGTTGCAATGAAAGAGGCTTTGTTGATTTGTTATTAACACAGAAATCAAAAGTTTC 360
Db |||||||
QY 382 AGAAAGAGTGTCTCAAGAAATGAGTGTCTCACTGGAATGCCGTCTTTGACCTGGCTGG 441
Db |||||||
QY 361 AGAAAGAGTGTCTCAAGAAATGAGTGTCTCACTGGAATGCCGTCTTTGACCTGGCTGG 420
Db |||||||
QY 442 GTTCTGCTGAGTAAACTTGTATACAGCAGCAGTGTATCAACAGCCAAATTTTGGGAC 501
Db |||||||
QY 421 GTTCTGCTGAGTAAACTTGTATACAGCAGCAGTGTATCAACAGCCAAATTTTGGGAC 480
Db |||||||
QY 502 GTAAAGCTGTGAGTGAATTTGAAACATGCAAGAGTCTCAATGAGCCTCAAGTCAGTT 561
Db |||||||
QY 481 GTAAAGCTGTGAGTGAATTTGAAACATGCAAGAGTCTCAATGAGCCTCAAGTCAGTT 540
Db |||||||
QY 562 GCCTTTCTAAGTTTGAAAGAGCTGTATTTCTGTACGGGTGGAAGAGATGGCAACATATG 621
Db |||||||
QY 541 GCCTTTCTAAGTTTGAAAGAGCTGTATTTCTGTACGGGTGGAAGAGATGGCAACATATG 600
Db |||||||
QY 622 GTCTGGGATACAGGTGCAACAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAGT 681
Db |||||||
QY 601 GTCTGGGATACAGGTGCAACAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAGT 660
Db |||||||
QY 682 GGAGCTCACAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGAAAGAGAAATCA 741
Db |||||||
QY 661 GGAGCTCACAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGAAAGAGAAATCA 720
Db |||||||
QY 742 AAAGGACTTGTCTCTTCTGTGGAATTTCCAGAAAGTGTATCTGTGGTCTCTTTCAAGAC 801
Db |||||||
QY 721 AAAGGACTTGTCTCTTCTGTGGAATTTCCAGAAAGTGTATCTGTGGTCTCTTTCAAGAC 780
Db |||||||
QY 802 GAGATACCTTAGTCTCAGCAGGAGCTGTGATGGGATTAATCAAAAGTATGGGATTTACGT 861
Db |||||||
QY 781 GAGATACCTTAGTCTCAGCAGGAGCTGTGATGGGATTAATCAAAAGTATGGGATTTACGT 840
Db |||||||
QY 862 AAGATTTATCTGCTTATCGACAAAGACCCATAGCATCCAAAGTCTTTCTGCTGACCCAGT 921
Db |||||||
QY 841 AAGATTTATCTGCTTATCGACAAAGACCCATAGCATCCAAAGTCTTTCTGCTGACCCAGT 900
Db |||||||
QY 922 AGCAGCACTCGAAACTTTGGATATTTCAAGTCTGATTTTGGATTTCCAGTGGCTCTACTTTA 981
Db |||||||
QY 901 AGCAGCACTCGAAACTTTGGATATTTCAAGTCTGATTTTGGATTTCCAGTGGCTCTACTTTA 960
Db |||||||
```

QY 982 TTTGCTAATTGCACAGACGATAACATCTACATGTTTAAATATGACTGGGTGGAAGCTTCT 1041
Db 961 TTTGCTAATTGCACAGACGATAACATCTACATGTTTAAATATGACTGGGTGGAAGCTTCT 1020
QY 1042 CCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT 1101
Db 1021 CCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT 1080
QY 1102 CCAGATGACCAAGTCTTTTAGTCAAGTGGCTCAAGTGAAGCTGCCCTACATATGGAAGGTC 1161
Db 1081 CCAGATGACCAAGTCTTTTAGTCAAGTGGCTCAAGTGAAGCTGCCCTACATATGGAAGGTC 1140
QY 1162 TCACACCTCGGACCTCTCTAATGCTGCTCCCTGGGTCAATCTCAAGAGTCAAGTCTGTG 1221
Db 1141 TCACACCTCGGACCTCTCTAATGCTGCTCCCTGGGTCAATCTCAAGAGTCAAGTCTGTG 1200
QY 1222 TGCTGGTGTCCATCTGACTTCAAAAGATTGCTACTGTTCTGATGACAATACACTAAAA 1281
Db 1201 TGCTGGTGTCCATCTGACTTCAAAAGATTGCTACTGTTCTGATGACAATACACTAAAA 1260
QY 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGAGGTGATAAACTTTCCACGGTG 1341
Db 1261 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGAGGTGATAAACTTTCCACGGTG 1320
QY 1342 GGTGGGCTCTCAGAGAAAAAGAGTCAAGACCTGGCTAGTAACTAAAGAGTAGC 1401
Db 1321 GGTGGGCTCTCAGAGAAAAAGAGTCAAGACCTGGCTAGTAACTAAAGAGTAGC 1380
QY 1402 CAGAGTACTCTCCGACCAAGCCCCAGGGTAAAGTGAATCCAACTTTTCCCGCTCA 1461
Db 1381 CAGAGTACTCTCCGACCAAGCCCCAGGGTAAAGTGAATCCAACTTTTCCCGCTCA 1440
QY 1462 TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCCTACG 1521
Db 1441 TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCCTACG 1500
QY 1522 TTCTCTATTAATAACCTCTCCTGCAAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTG 1581
Db 1501 TTCTCTATTAATAACCTCTCCTGCAAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTG 1560
QY 1582 TCCTCGCTCTCCCAAGCCACTTCACTCTTCAAGATGTCGATTAGAACTGGGTGACC 1641
Db 1561 TCCTCGCTCTCCCAAGCCACTTCACTCTTCAAGATGTCGATTAGAACTGGGTGACC 1620
QY 1642 CGAACACCTTCTCTCATCACACCCATCACTCCACCTGCTTCGGAGACCAAGATCATGCT 1701
Db 1621 CGAACACCTTCTCTCATCACACCCATCACTCCACCTGCTTCGGAGACCAAGATCATGCT 1680
QY 1702 CCGAGAAAAGCCCTTATTCCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGCTTGTCTGAG 1761
Db 1681 CCGAGAAAAGCCCTTATTCCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGCTTGTCTGAG 1740
QY 1762 TCTAGAAATAGATTAAGAGGAGGCTTAGACTCAAGCTGTCTGGAGGCTGGAACAAAG 1821
Db 1741 TCTAGAAATAGATTAAGAGGAGGCTTAGACTCAAGCTGTCTGGAGGCTGGAACAAAG 1800
QY 1822 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAATCTTCATTG 1881
Db 1801 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAATCTTCATTG 1860
QY 1882 GATCTGTGCTGCTGTGCTGTTAAACGAGACCTTTAGTAAGGACTCTCTAGGCTCTTACC 1941
Db 1861 GATCTGTGCTGCTGTGCTGTTAAACGAGACCTTTAGTAAGGACTCTCTAGGCTCTTACC 1920
QY 1942 AAATCAAGCAAAATTTGAAGGAGCTGTACAGTATCTCAGAGCCTCCGCTCCTATCAGT 2001
Db 1921 AAATCAAGCAAAATTTGAAGGAGCTGTGTACAGTATCTCAGAGCCTCCGCTCCTATCAGT 1980
QY 2002 CCGTATGCTTCAGAAAGCTGTGGAAGCTTACCTCTTCTTTGAGACCTTTGTGGAGAGGG 2061
Db 1981 CCGTATGCTTCAGAAAGCTGTGGAAGCTTACCTCTTCTTTGAGACCTTTGTGGAGAGGG 2040

QY 2062 TCTCAATGGTAGGCAAGAGATAATTTCCCGCAGAGATAAAACTGGTTGTTGGCCATG 2121
Db 2041 TCTCAATGGTAGGCAAGAGATAATTTCCCGCAGAGATAAAACTGGTTGTTGGCCATG 2100
QY 2122 GCAGCCAAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCGTCATCCAGACACCCCAAT 2181
Db 2101 GCAGCCAAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCGTCATCCAGACACCCCAAT 2160
QY 2182 TCCAGGAGACAGAGCGGAAGACATTCGCAAGCCCGGTCAACATCAAGCCAGCTCCATG 2241
Db 2161 TCCAGGAGACAGAGCGGAAGACATTCGCAAGCCCGGTCAACATCAAGCCAGCTCCATG 2220
QY 2242 AGGAAATCTGCACATATCTTCCATAGAAAGTCCAGAGGAGACTTCTGTGGTCCCTGAAAC 2301
Db 2221 AGGAAATCTGCACATATCTTCCATAGAAAGTCCAGAGGAGACTTCTGTGGTCCCTGAAAC 2280
QY 2302 TCAACAGAAATATATAGATTCTAATCTGAGTGAATTTACTGAGCTTTGGTCCCACTAAAAA 2361
Db 2281 TCAACAGAAATATATAGATTCTAATCTGAGTGAATTTACTGAGCTTTGGTCCCACTAAAAA 2340
QY 2362 CTGAGCTTTGGTCCCACTAAAAAAGATGAAAAATCAAGAGTGAATCTATTAATCTCTGTC 2421
Db 2341 CTGAGCTTTGGTCCCACTAAAAAAGATGAAAAATCAAGAGTGAATCTATTAATCTCTGTC 2400
QY 2422 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2481
Db 2401 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2460
QY 2482 TCTGCTTCTACTACCAATAATGTATATGAGCTTCCCGAGGATGAATGCTGTTTAAAT 2541
Db 2461 TCTGCTTCTACTACCAATAATGTATATGAGCTTCCCGAGGATGAATGCTGTTTAAAT 2520
QY 2542 TCATAAAGTAAATTTGTACCTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2601
Db 2521 TCATAAAGTAAATTTGTACCTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2580
QY 2602 TCATCTTCTCTAATAATGACATCCAGTTCATGGAGGCAAAAACAAGTTTCTTGTGA 2661
Db 2581 TCATCTTCTCTAATAATGACATCCAGTTCATGGAGGCAAAAACAAGTTTCTTGTGA 2640
QY 2662 TCCTGAAACTTTTCTATGCTCAGTGAAGATCTGCCAGCCACAGCATGAGGCCCTGTGAA 2721
Db 2641 TCCTGAAACTTTTCTATGCTCAGTGAAGATCTGCCAGCCACAGCATGAGGCCCTGTGAA 2700
QY 2722 GGCTGACTGAGAAATTCCTCTGCTGAAGACCCCTGGTTCTGTTCTGCTCCAAATGTATA 2781
Db 2701 GGCTGACTGAGAAATTCCTCTGCTGAAGACCCCTGGTTCTGTTCTGCTCCAAATGTATA 2760
QY 2782 ATTTTATTTGAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAATACATAATCTTTTCACTATG 2793

RESULT 3

US-11-266-748A-56960
; Sequence 56960, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnstons, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266.748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0

|||||
706 CAAAGTATGGATTATAGTAAGAAATATATCTGCTTATCGACAAGAACCCATAGCATCCAA 765
|||||
903 GTCTTTCTGTACCCAGGTAGCAGCACTCGAANAACCTTGGATATTCAAGTCTGATTTTGA 962
|||||
766 GTCTTTCTGTGATCCCAAGGTAGCAGCACTCGAANAACCTTGGATATTCAAGTCTGATTTTGA 925
|||||
963 TTCCACTGGCTCTPACTTATTTTCTAAATTGTCACAGACGATAACATCTACATGTTTAAATAT 1022
|||||
826 TTCCACTGGCTCTPACTTATTTGCTTAATTGTCACAGAGATACATCTACATGTTTAAATAT 885
|||||
1023 GACTGGGTTGAAGA CTCTCCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTA 1082
|||||
886 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACACAACTCTACCTTTTA 945
|||||
1083 TGTAAATCCAGCTTAGTCCAGATGACACAGATTTTGTAGTCAGTGGCTCAAGTGAAGAAG 1142
|||||
946 TGTAAATCCAGCTTAGTCCAGATGACACAGATTTTGTAGTCAGTGGCTCAAGTGAAGAAG 1005
|||||
1143 TGCTTACATATGGAAGGTCTCCACACCTCGGCAACCTCTACTGTGCTCTCTGGGTCAATTC 1202
|||||
1006 TGCTTACATATGGAAGGTCTCCACACCTCGGCAACCTCTACTGTGCTCTCTGGGTCAATTC 1065
|||||
1203 TCAAGAGGTCACTCTGTGTGTGTGTGTCCATCTGACTTCAAAAAGATTTGCTACTCTGTTTC 1262
|||||
1066 TCAAGAGGTCACTCTGTGTGTGTGTGTCCATCTGACTTCAAAAAGATTTGCTACTCTGTTTC 1125
|||||
1263 TGAATGACAATACACTAAAAATCTGGGGCTTGAATAGAGGCTTAGAGGAGAAACCCAGGAGG 1322
|||||
1126 TGAATGACAATACACTAAAAATCTGGGGCTTGAATAGAGGCTTAGAGGAGAAACCCAGGAGG 1185
|||||
1323 TGATAAATCTTCCACGGTGGTGGGCTCTCAGAGAAGAAAAGAGTCAAGACCTGGGCT 1382
|||||
1186 TGATAAATCTTCCACGGTGGTGGGCTCTCAGAGAAGAAAAGAGTCAAGACCTGGGCT 1245
|||||
1383 AGTAAACAGTAACAGTAGCAGAGTACTCTGTGCCAAAGCCCCCAGAGGTAAAGTGCAATCC 1442
|||||
1246 AGTAAACAGTAACAGTAGCAGAGTACTCTGTGCCAAAGCCCCCAGAGGTAAAGTGCAATCC 1305
|||||
1443 ATCCAAATCTTCCCGCTCATFCGACAGTGTGTGCCCAAGTGTGTGAGAGACTCCCTCT 1502
|||||
1306 ATCCAAATCTTCCCGCTCATFCGACAGTGTGTGCCCAAGTGTGTGAGAGACTCCCTCT 1365
|||||
1503 TCCTTCAATATCTCCATCTCTCTATTAATAAAGCTCTCTGCAAGGCGGCTCTCCAT 1562
|||||
1366 TCCTTCAATATCTCCATCTCTCTATTAATAAAGCTCTCTGCAAGGCGGCTCTCCAT 1425
|||||
1563 CAAAGAGAGGCTCTGTCTCTCTCTCTCTCCCAAGCCACCTTCTATCTTTTCAAGATGTC 1622
|||||
1426 CAAAGAGAGGCTCTGTCTCTCTCTCTCTCTCCCAAGCCACCTTCTATCTTTCAAGATGTC 1485
|||||
1623 GATTAGAAATCTGGGTGACCGGAACACCTTCTCTCTCATCAACCCATCACTCCACCTGCTTC 1682
|||||
1486 GATTAGAAATCTGGGTGACCGGAACACCTTCTCTCATCAACCCATCACTCCACCTGCTTC 1545
|||||
1683 GGAGACCAAGATCATCTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCA 1742
|||||
1546 GGAGACCAAGATCATCTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCA 1605
|||||
1743 AGCAGAGGCTTGTCTGTAGTCTAGAAATAGATAAGAGAGGCTTAGACTCAAGCTGTCT 1802
|||||
1606 AGCAGAGGCTTGTCTGTAGTCTAGAAATAGAGTAAAGAGAGGCTTAGACTCAAGCTGTCT 1665
|||||
1803 GGAGAGTGTGAACAAAGAGTGTGAAGAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAG 1862
|||||
1666 GGAGAGTGTGAACAAAGAGTGTGAAGAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAG 1725
|||||
1863 AGTTGAATCTTCAATTTGGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1922
|||||
1726 AGTTGAATCTTCAATTTGGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1785
|||||
1923 GGACTCTCTAGTCTTCAACCAATCAAGACAAATTTGAAGGAGCTGTGTACAGATATCTCACA 1982
|||||

Db 1786 GGACTCTCTAGTCTCTACCAAAATCAAGCAAAATTTGAAGGAGCTGTGTACCAAGTATCTCAGA 1845
Qy 1983 GCCTCCGCTCTCTATCAGTCCGTATGCTTTCAGAAAGCTGTGGAACGCTACCTCTTCCCTTT 2042
Db 1846 GCCTCCGCTCTCTATCAGTCCGTATGCTTTCAGAAAGCTGTGGAACGCTACCTCTTCCCTTT 1905
Qy 2043 GAGACTCTGTGTGAGAGGGTCTGAAATGTTAGGCAAGAGAAATAGTTTCCCAAGAAATAA 2102
Db 1906 GAGACTCTGTGTGAGAGGGTCTGAAATGTTAGGCAAGAGAAATAGTTTCCCAAGAAATAA 1965
Qy 2103 AAACCTGGTGTGGCCATGCGCAGCCAAACGGAAGCTGAGAAATCCATCTCAGCAAGTCC 2162
Db 1966 AAACCTGGTGTGGCCATGCGCAGCCAAACGGAAGCTGAGAAATCCATCTCAGCAAGTCC 2025
Qy 2163 GTTCATCCAGACACCAATTTCCAGGAGACAGAGCGGAAGACATTTCCCAAGCCCGTCCAC 2222
Db 2026 GTTCATCCAGACACCAATTTCCAGGAGACAGAGCGGAAGAAATTTCCCAAGCCCGTCCAC 2085
Qy 2223 CATCAGCCCGAGCTCCATGAGGAAATCTGCACATATTTCCA TAGAAAAGTCCCAAGGAGA 2282
Db 2086 CATCAGCCCGAGCTCCATGAGGAAATCTGCACATATTTCCA TAGAAAAGTCCCAAGGAGA 2145
Qy 2283 CTTCTGTGGTCTGAACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTACTGAGC 2342
Db 2146 CTTCTGTGGTCTGAACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTACTGAGC 2205
Qy 2343 TTTGGTCCACATAAACAAGCTGAGCTTTGGTCCACTAATAAACAAGATGAAAAATACAAGAG 2402
Db 2206 TTTGGTCCACATAAACAAGCTGAGCTTTGGTCCACTAATAAACAAGATGAAAAATACAAGAG 2265
Qy 2403 TGACTCTATACTCTGGTCTTTAAGAAAGCTGCCCTTTTCAATTTTATAGACAAAATCTTTTC 2462
Db 2266 TGACTCTATACTCTGGTCTTTAAGAAAGCTGCCCTTTTCAATTTTATAGACAAAATCTTTTC 2325
Qy 2463 AACGTGAAATGTAACCTAATCTGTTCTACTACCAATAATATATGCAAGTTCCTCCGAGGA 2522
Db 2326 AACGTGAAATGTAACCTAATCTGTTCTACTACCAATAATATATGCAAGTTCCTCCGAGGA 2385
Qy 2523 TGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGCTACTCTAGCATTTTGAATGAATAG 2582
Db 2386 TGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGCTACTCTAGCATTTTGAATGAATAG 2445
Qy 2583 TCTTCACCTTTTAAATTTTCAATCTCTCTATAATAATGAATCCAGTTCATGAGGCA 2642
Db 2446 TCTTCACCTTTTAAATTTTCAATCTCTCTATAATAATGAATCCAGTTCATGAGGCA 2505
Qy 2643 AAAAAAAGTTCCTGTTATCTGAAACTTTCTATGCTCAGTGGAAGTATCTGCCAGCC 2702
Db 2506 AAAAAAAGTTCCTGTTATCTGAAACTTTCTATGCTCAGTGGAAGTATCTGCCAGCC 2565
Qy 2703 ACAGCATGAGGCTGTGAGGCTGAGAAATCTCTGCTGAAGACCCCTGTTCTGT 2762
Db 2566 ACAGCATGAGGCTGTGAGGCTGAGAAATCTCTGCTGAAGACCCCTGTTCTGT 2625
Qy 2763 TCTGCTCTCAACATGATATAATTTTAAATAATATATTTTGAATAATATCTTTTCACTATG 2814
Db 2626 TCTGCTCTCAACATGATATAATTTTAAATAATATATTTTGAATAATATCTTTTCACTATG 2677

RESULT 5

US-11-266-748A-184890
; Sequence 184890, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2

```
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 184890
; LENGTH: 1949
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-184890

Query Match      68.4%; Score 1937.4; DB 8; Length 1949;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1941; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 12 GAGTTGGAGGCGATACGATTTGTGTGTGAGAGGCGCAACGTCGGGATTTCTGCTGAAC 71
DB 3 GGGTTGGAGGCGATACGATTTGTGTGTGAGAGGCGCAACGTCGGGATTTCTGCTGAAC 62

QY 72 TGGAGGCAATTTACGACATTTTCTCTCAGTCAGGCTTTTCTCCGACCTGATGCTCTT 131
DB 63 TGGAGGCAATTTCTACGACATTTTCTCTCAGTCAGGCTTTTCTCCGACCTGATGCTCTT 122

QY 132 CAATTCGGTGCTCCGCGACCCAGCTTCAGTCAGGCTTCCTGAGAAATGGATGCTTTCAACATA 191
DB 123 CAATTCGGTGCTCCGCGACCCAGCTTCAGTCAGGCTTCCTGAGAAATGGATGCTTTCAACATA 182

QY 192 CCCTCTTCAATCCCTTCTGACTGGTTATCAGTCGAGTGGTAAATGATGAACACATCTTTTA 251
DB 183 CCCTCTTCAATCCCTTCTGACTGGTTATCAGTCGAGTGGTAAATGATGAACACATCTTTTA 242

QY 252 TGGAGAAACAGAGTCCAGTTCCTCTTTGGATGTACCTTCTCTCTGCTCCCAATAT 311
DB 243 TGGAGAAACAGAGTCCAGTTCCTCTTTGGATGTACCTTCTCTCTGCTCCCAATAT 302

QY 312 GGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTGCAATGTATATAACACAGAATC 371
DB 303 GGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTGCAATGTATATAACACAGAATC 362

QY 372 ACAAGTTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGCTCACTGGAATGCCGTCTTTGA 431
DB 363 ACNAGTTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGCTCACTGGAATGCCGTCTTTGA 422

QY 432 CTGGCCTCGGTTCTCTGGTGAATTTAAACTTGTATACAGCAGGCTGATCAAAACAGCCAA 491
DB 423 CTGGCCTCGGTTCTCTGGTGAATTTAAACTTGTATACAGCAGGCTGATCAAAACAGCCAA 482

QY 492 ATTTTGGGACGTAAAGCTGGTGAGCTGATTTGMAATCATCAAGTTCATCAATGCAGCT 551
DB 483 ATTTTGGGACGTAAAGCTGGTGAGCTGATTTGMAATCATCAAGTTCATCAATGCAGCT 542

QY 552 CAAGTCAGTTGCTTTTCTTAAGTTTGGAGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGG 611
DB 543 CAAGTCAGTTGCTTTTCTTAAGTTTGGAGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGG 602

QY 612 CAACATTATGCTGCGGATACCAAGTGGTGAACAAAAAGATGGGTTTATAGGCAAGTGAA 671
DB 603 CAACATTATGCTGCGGATACCAAGTGGTGAACAAAAAGATGGGTTTATAGGCAAGTGAA 662

QY 672 TCAATACAGTGGGCTCAACATACCTCAGCAAGCAAAACCCCTTCAAAACCCCAAGAGAA 731
DB 672 TCAATACAGTGGGCTCAACATACCTCAGCAAGCAAAACCCCTTCAAAACCCCAAGAGAA 731
```

```
DB 663 TCAAAATCAGTGGAGCTCACAATACTCAGACAAGCAAAACCCCTTCAAAACCCCAAGAGAA 722
QY 732 ACAGAAATTCAAAAGAGCTTGTCTCTTCTGTGGATTTTCAGCAAAAGTGTACTGTGGTCT 791
DB 723 ACAGAAATTCAAAAGAGCTTGTCTCTTCTGTGGATTTTCAGCAAAAGTGTACTGTGGTCT 782
QY 792 CTTTTCAGAGACGAGAATACCTTAGTCTCAGCAGGAGCTGTGTGATGGGATAAATCAAAAGTATG 851
DB 783 CTTTTCAGAGACGAGAATACCTTAGTCTCAGCAGGAGCTGTGTGATGGGATAAATCAAAAGTATG 842
QY 852 GGATTTACGTAAGAATTATATCTGCTTATTCGAACAAGAACCCATAGCATCAAAGTCTTTTCT 911
DB 843 GGATTTACGTAAGAATTATATCTGCTTATTCGAACAAGAACCCATAGCATCAAAGTCTTTTCT 902
QY 912 GTACCCAGGTAGCAGACTCGAAAACCTTGATATTTCAAGTCTGATTTTGGATTCACATGG 971
DB 903 GTACCCAGGTAGCAGACTCGAAAACCTTGATATTTCAAGTCTGATTTTGGATTCACATGG 962
QY 972 CTCTACTTTATTTGCTAATTCGACAGAGATAAATCTACATGTTTATATATGACTGGGTT 1031
DB 963 CTCTACTTTATTTGCTAATTCGACAGAGATAAATCTACATGTTTATATATGACTGGGTT 1022
QY 1032 GAAGACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAATAATC 1091
DB 1023 GAAGACTTCTCCAGTGGCTATTTTCAATGGTGGCCAGAACTCTACCTTTTATGTAATAATC 1082
QY 1092 CAGCCTTAGTCCAGATGACAGTCTTTTAGTCAAGTGGCTCAAGTGAATGAAGTGCCTCAT 1151
DB 1083 CAGCCTTAGTCCAGATGACAGTCTTTTAGTCAAGTGGCTCAAGTGAATGAAGTGCCTCAT 1142
QY 1152 ATGGAAGTCTCCACACCTCGCAACCTCTTACTGTGCTCTGCTGGTCAATCTCAAGAGGT 1211
DB 1143 ATGGAAGTCTCCACACCTCGCAACCTCTTACTGTGCTCTGCTGGTCAATCTCAAGAGGT 1202
QY 1212 CACGTCGTGTGCTGCTGCATCTGACTTCAAAAGATTGCTACTGCTTCTGATGACAA 1271
DB 1203 CACGTCGTGTGCTGCTGCATCTGACTTCAAAAGATTGCTACTGCTTCTGATGACAA 1262
QY 1272 TACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGAGAGAAACAGAGAGGTGATAAATC 1331
DB 1263 TACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGAGAGAAACAGAGAGGTGATAAATC 1322
QY 1332 TTCCACGCTGGGTTGGGCTCTCAGAGAGAAAGAGTCAAGACTGSCCTAGTAACAGT 1391
DB 1323 TTCCACGCTGGGTTGGGCTCTCAGAGAGAAAGAGTCAAGACTGSCCTAGTAACAGT 1382
QY 1392 AACGAGTAGCCAGAGTACTCTCTGCAAAAGCCCCAGGGTAAAGTGCAATCCATCCAAATC 1451
DB 1383 AACGAGTAGCCAGAGTACTCTCTGCAAAAGCCCCAGGGTAAAGTGCAATCCATCCAAATC 1442
QY 1452 TTCCCGGTCATCCGAGCTTGTGCCCCCAAGCTGTGTGGAGACCTTCCCTCTTCTTCAAA 1511
DB 1443 TTCCCGGTCATCCGAGCTTGTGCCCCCAAGCTGTGTGGAGACCTTCCCTCTTCTTCAAA 1502
QY 1512 TACTCTCAGGTTCTCTATTAACCTCTCTGCAAGCCCGGCTCTCCCATCAACAGAG 1571
DB 1503 TACTCTCAGGTTCTCTATTAACCTCTCTGCAAGCCCGGCTCTCCCATCAACAGAG 1562
QY 1572 AGGCTCTGCTCTCTCCGCTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGATTTAGAAA 1631
DB 1563 AGGCTCTGCTCTCTCCGCTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGATTTAGAAA 1622
QY 1632 CTGGGTGACCCGAAACACCTTCTCTCATCAACCCATCACTCACTGCTTGGAGACCA 1691
DB 1623 CTGGGTGACCCGAAACACCTTCTCTCATCAACCCATCACTCACTGCTTGGAGACCA 1682
QY 1692 GATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCACTCCCAAGCAGAGGC 1751
DB 1683 GATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCACTCCCAAGCAGAGGC 1742
QY 1752 TTGCTCTGAGTCTAGAAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGT 1811
DB 1743 TTGCTCTGAGTCTAGAAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGT 1802
```


Query Match	43.6%;	Score 1235;	DB 8;	Length 1270;
Best Local Similarity	99.9%;	Prod. No. 0;		
Matches 1246;	Conservative	0; Mismatches	0; Indels	1; Gaps
1;				
QY	1202	CTCAAGAGGTCACTGCTGTGTGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTT	1261	
DB	24	CTCAAGAGGTCACTGCTGTGTGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTT	83	
QY	1262	CTGATGACAATACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCGAGG	1321	
DB	84	CTGATGACAATACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCGAGG	143	
QY	1322	GTGATTAACCTTCCACGGTGGGTGGGCGCTCTCAAGAGAAAAGAGTCAAGACCTCGCC	1381	
DB	144	GTGATTAACCTTCCACGGTGGGTGGGCGCTCTCAAGAGAAAAGAGTCAAGACCTCGCC	203	
QY	1382	TAGTAACAGTAACAGTAGTACGAGTACTCTCTGCCAAAGCCCCCAGGGTAAAGTGCAATC	1441	
DB	204	TAGTAACAGTAACAGTAGTACGAGTACTCTCTGCCAAAGCCCCCAGGGTAAAGTGCAATC	263	
QY	1442	CATCAAAATCTTCCCGGTCAATCCGAGCTGTGTGCCCAAGCTGTGTGGAGACCTTCCCTC	1501	
DB	264	CATCAAAATCTTCCCGGTCAATCCGAGCTGTGTGCCCAAGCTGTGTGGAGACCTTCCCTC	323	
QY	1502	TTCTCTTCAAATACTCTCTAGTCTCTATTAAACCTCTCTGCCAAAGCCCCCGGTCTCCCA	1561	
DB	324	TTCTCTTCAAATACTCTCTAGTCTCTATTAAACCTCTCTGCCAAAGCCCCCGGTCTCCCA	383	
QY	1562	TCAACAGAGAGGCTCTGTCTCTCCGTCTCTCCCAAGCCACCTTCATCTTTTCAAGATGT	1621	
DB	384	TCAACAGAGAGGCTCTGTCTCTCCGTCTCTCCCAAGCCACCTTCATCTTTTCAAGATGT	443	
QY	1622	CGATTAGAACTGGGTGACCCGAAACACCTTCTCATCACACCCATCATCTCCACCTGTT	1681	
DB	444	CGATTAGAACTGGGTGACCCGAAACACCTTCTCATCACACCCATCATCTCCACCTGTT	503	
QY	1682	CGGAGACCAAGATCATGTCTCCGAGAAAAGCCCTTATTCTGTGAGCCAGAAAGTCATCCC	1741	
DB	504	CGGAGACCAAGATCATGTCTCCGAGAAAAGCCCTTATTCTGTGAGCCAGAAAGTCATCCC	563	
QY	1742	AAGCAGAGCTTGTCTCTGAGTCTAGAAATAGAGTAAAGAGAGGCTTAGACTCAAGCTGTC	1801	
DB	564	AAGCAGAGCTTGTCTCTGAGTCTAGAAATAGAGTAAAGAGAGGCTTAGACTCAAGCTGTC	623	
QY	1802	TCGAGAGTGTAAACAAAGTGTGTGAAGGTGTGTAACCTGTGTGACTGAGCTTCATGCCC	1861	
DB	624	TCGAGAGTGTAAACAAAGTGTGTGAAGGTGTGTAACCTGTGTGACTGAGCTTCATGCCC	683	
QY	1862	AAAGTTGAAAATCTTCAATTTGATCTGTGCTGCTTGTGTTGCTGGTAAACCAAGAGACCTTAGTA	1921	

Db 768 AAATCCTCTGCTGAAGACCCCTGGTTCTGTTCTGCTCCCAACATGTATTAATTTTGA 827
QY 2793 AATACATAATCTTTTCACTATG 2814
Db 828 AATACATAATCTTTTCACTATG 849

RESULT 11
US-11-266-748A-146319/c
; Sequence 146319, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 146319
; LENGTH: 931
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (62)..(71)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-146319

Query Match 28.3%; Score 802; DB 8; Length 931;
Best Local Similarity 100.0%; Pred.No. 1.3e-213; Indels 0; Gaps 0;
Matches 802; Conservative 0; Mismatches 0;
QY 2013 AGAAGCTGTGGAACGCTACCTCTTCTTTGAGACCTCTGTGGAGAGGCTGTGAATGT 2072
Db 884 AGAAGCTGTGGAACGCTACCTCTTCTTTGAGACCTCTGTGGAGAGGCTGTGAATGT 825
QY 2073 AGGCAAGAGATAGTTCCTCCAGAGATATAAATCTGGTTGTGGCCATGGCAGCAACG 2132
Db 824 AGGCAAGAGATAGTTCCTCCAGAGATATAAATCTGGTTGTGGCCATGGCAGCAACG 765
QY 2133 GAAGGCTGAGATCCATCTCCAGAGTCCGTCATCCAGACACCCCAATTCAGGAGACA 2192
Db 764 GAAGGCTGAGATCCATCTCCAGAGTCCGTCATCCAGACACCCCAATTCAGGAGACA 705
QY 2193 GAGCGAAGACATTCGCCAAGCCCGGTCCACCATCAGCCCGAGCTCCATCAGGAAATCTG 2252
Db 704 GAGCGAAGACATTCGCCAAGCCCGGTCCACCATCAGCCCGAGCTCCATCAGGAAATCTG 645
QY 2253 CACATATCTCCATAGAAAGTCCAGAGGACTTCTGTGTCTCTGACACTCAACAGAAAT 2312
Db 644 CACATATCTCCATAGAAAGTCCAGAGGACTTCTGTGTCTCTGACACTCAACAGAAAT 585

QY 2313 ATAGATTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACCAAGCTGAGCTTTGG 2372
Db 584 ATAGATTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACCAAGCTGAGCTTTGG 525
QY 2373 TCCACTAAACCAAGATGAAAAATACAAAGTGACTCTATACTCTGGTCTTTAAGAAAGC 2432
Db 524 TCCACTAAACCAAGATGAAAAATACAAAGTGACTCTATACTCTGGTCTTTAAGAAAGC 465
QY 2433 TGGCTTTTCAATTTTAGACAAATCTTTCAACGCTGAAATGTAATCTATCTGTTCTAC 2492
Db 464 TGGCTTTTCAATTTTAGACAAATCTTTTCAACGCTGAAATGTAATCTATCTGTTCTAC 405
QY 2493 TACCATAATGTATATGCACTTCCGAGAGTGAATGCTGTGTTAAATTTCAAAAGTAA 2552
Db 404 TACCATAATGTATATGCACTTCCGAGAGTGAATGCTGTGTTAAATTTCAAAAGTAA 345
QY 2553 ATTTCTCACTCTAGCATTTTGAATGAATGACTTCTCACTTTTAAATTTAATCTTCTCT 2612
Db 344 ATTTCTCACTCTAGCATTTTGAATGAATGACTTCTCACTTTTAAATTTAATCTTCTCT 285
QY 2613 ATAATAATGACATCCAGTTCATGAGGCAAAACAAAGTTTCTGTTATCTCTGAAACTT 2672
Db 284 ATAATAATGACATCCAGTTCATGAGGCAAAACAAAGTTTCTGTTATCTCTGAAACTT 225
QY 2673 TCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAAGGCTGACTGAG 2732
Db 224 TCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAAGGCTGACTGAG 165
QY 2733 AAATCCTCTGCTGAAGACCCCTGTTCTGTTCTGCTCCACATGTATATAATTTATTTGA 2792
Db 164 AAATCCTCTGCTGAAGACCCCTGTTCTGTTCTGCTCCACATGTATATAATTTATTTGA 105
QY 2793 AATACATAATCTTTTCACTATG 2814
Db 104 AATACATAATCTTTTCACTATG 83

RESULT 12
US-11-266-748A-41156
; Sequence 41156, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 41156
; LENGTH: 883
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-41156

```
Query Match      28.2%; Score 799.4; DB 8; Length 883;
Best Local Similarity 99.9%; Pred. No. 6.9e-213;
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2014 GAAAGCTGTGGAAGCTACCTCTCTCTTTGAGACCTTGTGAGAGGGTCTGAAATGTA 2073
DB 1 GAAAGCTGTGGAAGCTACCTCTCTCTTTGAGACCTTGTGAGAGGGTCTGAAATGTA 60

QY 2074 GGCAAGAGAGATAGTTCCTCCAGAGAAATAAAACCTGTTGGCCATGGCAGCCAAACGG 2133
DB 61 GGCAAGAGAGATAGTTCCTCCAGAGAAATAAAACCTGTTGGCCATGGCAGCCAAACGG 120

QY 2134 AAGGCTGAGATCCATCTCCAGAGATCCGTCATCCAGACACCCCAATTCAGGAGACAG 2193
DB 121 AAGGCTGAGATCCATCTCCAGAGATCCGTCATCCAGACACCCCAATTCAGGAGACAG 180

QY 2194 AGCGAAAGACATTCGCAAGCCCGTCCACATCAGCCCGCTCCATGAGGAAATCTGC 2253
DB 181 AGCGAAAGACATTCGCAAGCCCGTCCACATCAGCCCGCTCCATGAGGAAATCTGC 240

QY 2254 ACATACCTTCCATAGAAAAGTCCAGGAGACTTCTGTGCTCCTGAACACTCAACAGAAATTA 2313
DB 241 ACATACCTTCCATAGAAAAGTCCAGGAGACTTCTGTGCTCCTGAACACTCAACAGAAATTA 300

QY 2314 TAGATTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAAACAGCTGAGCTTTGGT 2373
DB 301 TAGATTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAAACAGCTGAGCTTTGGT 360

QY 2374 CCACCTAAAACAAGATGAAAATAACAAGATGACTCTATAACTCTGTTTAAAGAAAGCT 2433
DB 361 CCACCTAAAACAAGATGAAAATAACAAGATGACTCTATAACTCTGTTTAAAGAAAGCT 420

QY 2434 GCCTTTTCAATTTTGTAGACAAAATCTTTTCAACGCTGAAAATGACTAATCTGGTCTACT 2493
DB 421 GCCTTTTCAATTTTGTAGACAAAATCTTTTCAACGCTGAAAATGACTAATCTGGTCTACT 480

QY 2494 ACCATAATGATATATGACAGCTTCCCGAGGATGAATGCTGTTTAAATTTCAATGAATAA 2553
DB 481 ACCATAATGATATATGACAGCTTCCCGAGGATGAATGCTGTTTAAATTTCAATGAATAA 540

QY 2554 TTTGTCACTCTAGCAATTTTGAATGAAATGCTTCACTTTTAAATTTTAAATTTCTCTCTA 2613
DB 541 TTTGTCACTCTAGCAATTTTGAATGAAATGCTTCACTTTTAAATTTTAAATTTCTCTCTA 600

QY 2614 TAATAATGACATCCAGTTTCATGAGGCAAAAACAAAGTTCTTGTGTTTCTGTTTCTGAAACTTT 2673
DB 601 TAATAATGACATCCAGTTTCATGAGGCAAAAACAAAGTTCTTGTGTTTCTGTTTCTGAAACTTT 660

QY 2674 CTATGCTCAGTGGAAAGTATCTGCCAGGCCACAGCATGAGGCTGTGAAGGCTGACTGAGA 2733
DB 661 CTATGCTCAGTGGAAAGTATCTGCCAGGCCACAGCATGAGGCTGTGAAGGCTGACTGAGA 720

QY 2734 AATCCTCTGCTGAAGACCCCTGGTTCTGTCTGCTCCCAACATGATATAATTTTTCGAA 2793
DB 721 AATCCTCTGCTGAAGACCCCTGGTTCTGTCTGCTCCCAACATGATATAATTTTTCGAA 780

QY 2794 ATACATAATCTTTTCACTATG 2814
DB 781 ATACATAATCTTTTCACTATG 801
```

RESULT 13
US-11-266-748A-210423
; Sequence 210423, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)

```
; CURRENT APPLICATION NUMBER: US/11/266,748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662,276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700,293  
; PRIOR FILING DATE: 2005-07-18  
; NUMBER OF SEQ ID NOS: 483996  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 210423  
; LENGTH: 883  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; US-11-266-748A-210423
```

```
Query Match      28.2%; Score 799.4; DB 8; Length 883;  
Best Local Similarity 99.9%; Pred. No. 6.9e-213;  
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2014 GAAAGCTGTGGAAGCTACCTCTCTCTTTGAGACCTTGTGAGAGGGTCTGAAATGTA 2073  
DB 1 GAAAGCTGTGGAAGCTACCTCTCTCTTTGAGACCTTGTGAGAGGGTCTGAAATGTA 60  
  
QY 2074 GGCAAGAGAGATAGTTCCTCCAGAGAAATAAAACCTGTTGGCCATGGCAGCCAAACGG 2133  
DB 61 GGCAAGAGAGATAGTTCCTCCAGAGAAATAAAACCTGTTGGCCATGGCAGCCAAACGG 120  
  
QY 2134 AAGGCTGAGATCCATCTCCAGAGATCCGTCATCCAGACACCCCAATTCAGGAGACAG 2193  
DB 121 AAGGCTGAGATCCATCTCCAGAGATCCGTCATCCAGACACCCCAATTCAGGAGACAG 180  
  
QY 2194 AGCGAAAGACATTCGCAAGCCCGTCCACATCAGCCCGCTCCATGAGGAAATCTGC 2253  
DB 181 AGCGAAAGACATTCGCAAGCCCGTCCACATCAGCCCGCTCCATGAGGAAATCTGC 240  
  
QY 2254 ACATACCTTCCATAGAAAAGTCCAGGAGACTTCTGTGCTCCTGAACACTCAACAGAAATTA 2313  
DB 241 ACATACCTTCCATAGAAAAGTCCAGGAGACTTCTGTGCTCCTGAACACTCAACAGAAATTA 300  
  
QY 2314 TAGATTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAAACAGCTGAGCTTTGGT 2373  
DB 301 TAGATTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAAACAGCTGAGCTTTGGT 360  
  
QY 2374 CCACCTAAAACAAGATGAAAATAACAAGATGACTCTATAACTCTGTTTAAAGAAAGCT 2433  
DB 361 CCACCTAAAACAAGATGAAAATAACAAGATGACTCTATAACTCTGTTTAAAGAAAGCT 420  
  
QY 2434 GCCTTTTCAATTTTGTAGACAAAATCTTTTCAACGCTGAAAATGACTAATCTGGTCTACT 2493  
DB 421 GCCTTTTCAATTTTGTAGACAAAATCTTTTCAACGCTGAAAATGACTAATCTGGTCTACT 480  
  
QY 2494 ACCATAATGATATATGACAGCTTCCCGAGGATGAATGCTGTTTAAATTTCAATGAATAA 2553  
DB 481 ACCATAATGATATATGACAGCTTCCCGAGGATGAATGCTGTTTAAATTTCAATGAATAA 540  
  
QY 2554 TTTGTCACTCTAGCAATTTTGAATGAAATGCTTCACTTTTAAATTTTAAATTTCTCTCTA 2613  
DB 541 TTTGTCACTCTAGCAATTTTGAATGAAATGCTTCACTTTTAAATTTTAAATTTCTCTCTA 600  
  
QY 2614 TAATAATGACATCCAGTTTCATGAGGCAAAAACAAAGTTCTTGTGTTTCTGTTTCTGAAACTTT 2673  
DB 601 TAATAATGACATCCAGTTTCATGAGGCAAAAACAAAGTTCTTGTGTTTCTGTTTCTGAAACTTT 660
```

Db	74	TTTACGTAAGAAATTATAC	TGCTGTTATGCAAGAAGAACCCATAGCATCCAAAGTCTTTCTCGTGA	133
Qy	915	CCCAGGTAGCAGCACTCG	AAACCTTGGATATTCAAAGTCTGATTTTGGATTTCCATCGCTC	974
Db	134	CCCAGGTAGCAGCACTCG	AAACCTTGGATATTCAAAGTCTGATTTTGGATTTCCATCGCTC	193
Qy	975	TACTTTTATTTGCTAA	TTTGGCAGAGATTAACATCTACATGTTTAAATGACTGGGTTGAA	1034
Db	194	TACTTTTATTTGCTAA	TTTGGCAGAGATTAACATCTACATGTTTAAATGACTGGGTTGAA	253
Qy	1035	GACTTCTCCAGTGGCTA	TTTTCAATGACACACCGAAGCTCTACCTTTTATGTAAATCCAG	1094
Db	254	GACTTCTCCAGTGGCTA	TTTTCAATGACACACCGAAGCTCTACCTTTTATGTAAATCCAG	313
Qy	1095	CTTTAGTCCAGATGACC	AGTTTATGTCAGTGGCTCAAAGTGTGAAGCTGCCTACATATG	1154
Db	314	CTTTAGTCCAGATGACC	AGTTTATGTCAGTGGCTCAAAGTGTGAAGCTGCCTACATATG	373
Qy	1155	GAAGGTCTCACACCTTG	GGCAACCTCTACTGTGCTCTCTGGGTCAATTTCTCAAGAGGTAC	1214
Db	374	GAAGGTCTCACACCTTG	GGCAACCTCTACTGTGCTCTCTGGGTCAATTTCTCAAGAGGTAC	433
Qy	1215	GTCTGTGTGCTGGTGC	ATCTGACATTCACAAAGATTTGCTACCTGTGTTCTCATGACAAATC	1274
Db	434	GTCTGTGTGCTGGTGC	ATCTGACATTCACAAAGATTTGCTACCTGTGTTCTCATGACAAATC	493
Qy	1275	ACTAAAAATCTGGCGC	TTGAATAGAGGCTTTAGAGGAGAAAAACAGGAGGTGATAACTTTTC	1334
Db	494	ACTAAAAATCTGGCGC	TTGAATAGAGGCTTTAGAGGAGAAAAACAGGAGGTGATAACTTTTC	553
Qy	1335	CACGGTGGGTTGGCCCT	CTCAGAGAAAAAGAGTCAAGACCTGGCCCTAGTAAACAGTAAAC	1394
Db	554	CACGGTGGGTTGGCCCT	CTCAGAGAAAAAGAGTCAAGACCTGGCCCTAGTAAACAGTAAAC	613
Qy	1395	-GAGTACGACAGTACTC	CTGCGCAAGACCCCA--GGTAAAGTGCATCCATCCCAATTC	1451
Db	614	TGATGACGACAGTACTC	CTGCGCAAGACCCCA--GGTAAAGTGCATCCATCCCAATTC	673
Qy	1452	TTCCCCGTCAATCCGAC	GTTGTGCCCCCAAGCTGTGCT---GGAGACCTCTCCCTCTTCCTT	1507
Db	674	TTCCCCGTCAATCCGAC	GTTGTGCCCCCAAGCTGTGCT---GGAGACCTCTCCCTCTTCCTT	733
Qy	1508	CAAAATCTCTTACG-TT	CTCTATTAAACCTCTCTGCGCAAGGCCCGGTCTCCCATCAA	1565
Db	734	CAAAATCTCTTACG-TT	CTCTATTAAACCTCTCTGCGCAAGGCCCGGTCTCCCATCAA	792
RESULT 15				
US-11-266-748A-322452/c				
; Sequence 322452, Application US/11266748A				
; Publication No. US20060134663A1				
; GENERAL INFORMATION:				
; APPLICANT: Harkin, Paul				
; APPLICANT: Johnston, Patrick				
; APPLICANT: Mulligan, Karl				
; TITLE OF INVENTION: Transcription Microarray Technology and				
; TITLE OF INVENTION: Methods of Using the Same				
; FILE REFERENCE: 55815-0102 (319189)				
; CURRENT APPLICATION NUMBER: US/11/266, 748A				
; CURRENT FILING DATE: 2005-11-03				
; PRIOR APPLICATION NUMBER: EP 04105479.2				
; PRIOR FILING DATE: 2004-11-03				
; PRIOR APPLICATION NUMBER: EP 04105482.6				
; PRIOR FILING DATE: 2004-11-03				
; PRIOR APPLICATION NUMBER: EP 04105483.4				
; PRIOR FILING DATE: 2004-11-03				
; PRIOR APPLICATION NUMBER: EP 04105507.0				
; PRIOR FILING DATE: 2004-11-03				
; PRIOR APPLICATION NUMBER: EP 04105485.9				
; PRIOR FILING DATE: 2004-11-03				
; PRIOR APPLICATION NUMBER: EP 04105484.2				
; PRIOR FILING DATE: 2004-11-03				
; PRIOR APPLICATION NUMBER: US 60/662,276				

;; PRIOR FILING DATE: 2005-03-14
;; PRIOR APPLICATION NUMBER: US 60/700,293
;; PRIOR FILING DATE: 2005-07-18
;; NUMBER OF SEQ ID NOS: 483996
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 322452
;; LENGTH: 798
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (39)..(39)
;; OTHER INFORMATION: n is a, c, g, or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (71)..(71)
;; OTHER INFORMATION: n is a, c, g, or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (93)..(93)
;; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-322452

Query Match 24.8%; Score 700.8; DB 8; Length 798;
Best Local Similarity 97.0%; Pred.No.2.5e-185;
Matches 756; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY 795 TCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGA 854
DB 785 TCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGA 726

QY 855 TTTACGTAAGAATTATATCTGCTTATCGACAAGAACCCATAGCATCCAAAGTCTTTCTCTGTA 914
DB 725 TTTACGTAAGAATTATATCTGCTTATCGACAAGAACCCATAGCATCCAAAGTCTTTCTCTGTA 666

QY 915 CCCAGGTACGACACTCGAAGAACTTTGGATTTCAAGTCTGATTTGGATTCCACTGGGCTC 974
DB 665 CCCAGGTACGACACTCGAAGAACTTTGGATTTCAAGTCTGATTTGGATTCCACTGGGCTC 606

QY 975 TACTTTATTTGCTAATTGACAGAGATAAACAATCTACATGTTTAAATATGACTGGGTTGAA 1034
DB 605 TACTTTATTTGCTAATTGACAGAGATAAACAATCTACATGTTTAAATATGACTGGGTTGAA 546

QY 1035 GACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAATCCAG 1094
DB 545 GACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAATCCAG 486

QY 1095 CCTTAGTCCAGATGACCACTTTTATGTCAGTGGCTCAAGTGATGAAGCTGCCCTCATATG 1154
DB 485 CCTTAGTCCAGATGACCACTTTTATGTCAGTGGCTCAAGTGATGAAGCTGCCCTCATATG 426

QY 1155 GAAGGTCTCCACACCTGGCAACCTCTACTGTCTGCTCTGGGTCAATTTCTCAAGAGGTAC 1214
DB 425 GAAGGTCTCCACACCTGGCAACCTCTACTGTCTGCTCTGGGTCAATTTCTCAAGAGGTAC 366

QY 1215 GTCTGTGTGCTGGTGTCCATCTGACTTCAAAAGATTTGCTACCTGTTCTGTATGACAATAC 1274
DB 365 GTCTGTGTGCTGGTGTCCATCTGACTTCAAAAGATTTGCTACCTGTTCTGTATGACAATAC 306

QY 1275 ACTAAAAATCTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACCAGAGGCTGATAAACTTTC 1334
DB 305 ACTAAAAATCTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACCAGAGGCTGATAAACTTTC 246

QY 1335 CACGGTGGGTTGGGCTCTCTCAGAAGAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAAC 1394
DB 245 CACGGTGGGTTGGGCTCTCTCAGAAGAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAAC 186

QY 1395 -GAGTAGCCAGAGTACTCTCGCCAAAGCCCCCA--GGGTAAAGTGCAATCCATCCAAATTC 1451
DB 185 TGAGTAGCCAGAGTACTCTCGCCAAAGCCCCCAAGCGGTAAATGTGCAATCCATCCAAATTC 126

QY 1452 TTCCCGCTCATCCGAGCTTTGTGCCCAAGCTGTCT---GGAGACCTCCCTCTTCCTT 1507
DB 1507 TTCCCGCTCATCCGAGCTTTGTGCCCAAGCTGTCT---GGAGACCTCCCTCTTCCTT

Db 125 TTCCCGCTCATCCGAGCTTTGTGCCCAAGCTTTGTGGACGACCTCCCTCTCTNTCCTT 66
QY 1508 CAAATATCTCCTACG-TTCTCTATTAAACCTCTCTGCGCAAGGCCCGGTCTCCCATCAA 1565
DB 65 CAAATATCTCCTACGTTTCTCTATTAAACCTCTCTGCGCAATGCCCGGTCTCCCATCAA 7

Search completed: November 9, 2006, 07:36:34
Job time : 645 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 06:20:06 ; Search time 4564 Seconds

(without alignments)
7621.888 Million cell updates/sec

Title: US-10-726-160-1

Perfect score: 2831

Sequence: 1 ggcacgagggggtggag.....atgaaaaa.....2831

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:

- 1: /EMC_Celerra_SID83/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /EMC_Celerra_SID83/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 3: /EMC_Celerra_SID83/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 4: /EMC_Celerra_SID83/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 5: /EMC_Celerra_SID83/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 6: /EMC_Celerra_SID83/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 7: /EMC_Celerra_SID83/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 8: /EMC_Celerra_SID83/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 9: /EMC_Celerra_SID83/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 10: /EMC_Celerra_SID83/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 11: /EMC_Celerra_SID83/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 12: /EMC_Celerra_SID83/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 13: /EMC_Celerra_SID83/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 14: /EMC_Celerra_SID83/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
- 15: /EMC_Celerra_SID83/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
- 16: /EMC_Celerra_SID83/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2831	100.0	2831	7	US-10-409-511-1
2	2831	100.0	2831	9	US-10-726-160-1
3	2816.2	99.5	2838	3	US-09-780-053-1
4	2804	99.0	4422	3	US-09-814-353-20186
5	2804	99.0	4422	9	US-10-357-930-25699
6	2789.8	98.5	4221	6	US-10-145-396-28
7	2789.8	98.5	4221	7	US-10-393-590-85
8	2789.8	98.5	4221	7	US-10-393-567-85
9	2789.8	98.5	4221	7	US-10-394-087-85
10	2789.8	98.5	4221	7	US-10-172-118-1645
11	2789.8	98.5	4221	8	US-10-342-887-1645
12	2789.8	98.5	4221	10	US-10-848-755A-1179
13	2442.4	86.3	4083	11	US-10-330-773-673
14	1520	53.7	1871	8	US-10-221-625-131
15	1302.2	46.0	2426	11	US-10-330-773-670
16	835.8	29.5	88892	11	US-10-330-773-672
17	649.8	23.0	57860	11	US-10-330-773-669

18	613	21.7	745	11	US-10-330-773-675	Sequence 675, App
19	558	19.7	708	6	US-10-076-555-715	Sequence 715, App
20	558	19.7	708	10	US-10-779-543-715	Sequence 715, App
21	507	17.9	507	6	US-10-206-901B-11	Sequence 11, Appl
22	500.4	17.7	815	10	US-10-779-543-3894	Sequence 3894, Ap
23	440	15.5	468	3	US-09-918-995-19957	Sequence 19957, A
24	439	15.5	543	3	US-09-814-353-16730	Sequence 16730, A
25	430.6	15.2	508	3	US-09-814-353-18083	Sequence 18083, A
26	421.4	14.9	445	3	US-09-780-053-3	Sequence 3, Appli
27	421.4	14.9	477	10	US-10-779-543-5959	Sequence 5959, Ap
28	411.6	14.5	434	3	US-09-814-353-5412	Sequence 5412, Ap
29	411.6	14.5	434	3	US-09-814-353-11699	Sequence 11699, A
30	403.4	14.2	493	6	US-10-106-698-491	Sequence 491, App
31	394.4	13.9	482	3	US-09-814-353-4038	Sequence 4038, Ap
32	394.4	13.9	482	3	US-09-814-353-10346	Sequence 10346, A
33	381.8	13.5	2001	8	US-10-403-571-73	Sequence 73, Appl
34	338	11.9	430	8	US-10-430-201-299	Sequence 299, App
35	338	11.9	430	8	US-10-430-201-300	Sequence 300, App
36	300	10.6	300	6	US-10-076-555-51	Sequence 51, Appl
37	300	10.6	300	10	US-10-779-543-51	Sequence 51, Appl
38	300	10.6	300	10	US-10-779-543-943	Sequence 943, App
39	300	10.6	300	10	US-10-779-543-1894	Sequence 1894, Ap
40	290.4	10.3	346	3	US-09-880-107-3634	Sequence 3634, Ap
41	282	10.0	297	9	US-10-357-930-8883	Sequence 8883, Ap
42	270	9.5	640	9	US-10-357-930-38800	Sequence 38800, A
43	254.4	9.0	432	16	US-11-128-061-1427	Sequence 1427, Ap
44	254.4	9.0	432	16	US-11-128-061-5069	Sequence 5069, Ap
45	254.4	9.0	432	16	US-11-128-049-1427	Sequence 1427, Ap

ALIGNMENTS

RESULT 1

US-10-409-511-1
; Sequence 1, Application US/10409511
; Publication No. US20030162268A1
; GENERAL INFORMATION:
; APPLICANT: IP, Nancy Y
; APPLICANT: CHEUNG, William M W
; TITLE OF INVENTION: No. US20030162268A1e1 Gene
; FILE REFERENCE: M98/0553/US
; CURRENT APPLICATION NUMBER: US/10/409,511
; CURRENT FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2831
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(2313)
; US-10-409-511-1

Query Match	100.0%;	Score 2831;	DB 7;	Length 2831;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2831;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	GGCACGAGCGGAGTTGGAGGCGATAACGATTGTGTGTGTGAGAGCGGCAACGTGCGATT	60	
Db	1	GGCACGAGCGGAGTTGGAGGCGATAACGATTGTGTGTGTGAGAGCGGCAACGTGCGATT	60	
Qy	61	TCGTCTGAATCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACC	120	
Db	61	TCGTCTGAATCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACC	120	
Qy	121	CTGATCTCTTCAATTCGGTGTCTCCGACGCCAGCTTGGCGCTCTGAGAAATGATGAG	180	
Db	121	CTGATCTCTTCAATTCGGTGTCTCCGACGCCAGCTTGGCGCTCTGAGAAATGATGAG	180	
Qy	181	TCFTTCAATACCTCTTCAATCCCTTCTGACTGTTTATCAGTGCAGTGGTAATGATGAA	240	
Db	181	TCFTTCAATACCTCTTCAATCCCTTCTGACTGTTTATCAGTGCAGTGGTAATGATGAA	240	

181 TCTTCACAATACCCCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAATGATGAA 240
241 CACACTTCTTATGAGAAACAGAGATCCAGATTCCTCTCTTTTGGATGATACCTTCTCTTCT 300
241 CACACTTCTTATGAGAAACAGAGATCCAGATTCCTCTCTTTTGGATGATACCTTCTCTTCT 300
301 GCTCCCAATATGGAACATGCTAGCAGTTGCCATGGAAGAGGCTTCTTGGATGGTAT 360
301 GCTCCCAATATGGAACATGCTAGCAGTTGCCAATGGAAGAGGCTTCTTGGATGGTAT 360
361 AACACAGAAATCAAAAGTTTCAGAAAGAGTGGCTTCAAAAGAAATGGATGGCTCACTGGAA 420
361 AACACAGAAATCAAAAGTTTCAGAAAGAGTGGCTTCAAAAGAAATGGATGGCTCACTGGAA 420
421 GCCGTCTTTGACCTGAGCTGGCTTCTGGTGAACCTTTAAATCTTGTACAGCAGCAGGTGAT 480
421 GCCGTCTTTGACCTGAGCTGGCTTCTGGTGAACCTTTAAATCTTGTACAGCAGCAGGTGAT 480
481 CAAACAGCCCAATTTTGGGACGTAAAGCTGGTGAAGCTGATTTGGAAATGCAAGGTGAT 540
481 CAAACAGCCCAATTTTGGGACGTAAAGCTGGTGAAGCTGATTTGGAAATGCAAGGTGAT 540
541 CAATGAGCTCAAGTCAGTTGGCTTTTCTAAGTTTGAAGAGCTGATTTCTGTACGGGT 600
541 CAATGAGCTCAAGTCAGTTGGCTTTTCTAAGTTTGAAGAGCTGATTTCTGTACGGGT 600
601 GGAAGAGATGGCAACATTAATGCTCTGGGATACAGGTGCAACAAAAGATGGGTTTAT 660
601 GGAAGAGATGGCAACATTAATGCTCTGGGATACAGGTGCAACAAAAGATGGGTTTAT 660
661 AGGCAAGTGAATCAATCAGTGGAGCTCAATACCTCAGACAGCAAGCAACCCCTTCAAAA 720
661 AGGCAAGTGAATCAATCAGTGGAGCTCAATACCTCAGACAGCAAGCAACCCCTTCAAAA 720
721 CCCAAGAAAGAACAGAAATCAAAAGACTTGTCTTCTGTGGATTTCCAGCAAAAGTGT 780
721 CCCAAGAAAGAACAGAAATCAAAAGACTTGTCTTCTGTGGATTTCCAGCAAAAGTGT 780
781 ACTGTGGTCTCTTTTCAAGACGAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
781 ACTGTGGTCTCTTTTCAAGACGAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
841 ATCAAAAGTATGGGATTTACGTAGAATATATCTGCTTATCGACAGAAACCCATAGCATCC 900
841 ATCAAAAGTATGGGATTTACGTAGAATATATCTGCTTATCGACAGAAACCCATAGCATCC 900
901 AAGTCTTTCTGTACCCAGGTAGCAGCTCGAAACCTTGGATTTCAAGTCTGATTTTG 960
901 AAGTCTTTCTGTACCCAGGTAGCAGCTCGAAACCTTGGATTTCAAGTCTGATTTTG 960
961 GATTCACCTGGCTCTACTTTATTTGCTAATTTGCACAGAGATACATCTACATGTTTAA 1020
961 GATTCACCTGGCTCTACTTTATTTGCTAATTTGCACAGAGATACATCTACATGTTTAA 1020
1021 ATGACTGGGTGGAAGCTTCTCAGTGGCTAATTTTCAATGGACACAGAACTCTACCTTT 1080
1021 ATGACTGGGTGGAAGCTTCTCAGTGGCTAATTTTCAATGGACACAGAACTCTACCTTT 1080
1081 TATGTAAAATCCAGCCTTAGTCCAGATGACCAAGTTTATAGTCAAGTGGCTCAAGTATGAA 1140
1081 TATGTAAAATCCAGCCTTAGTCCAGATGACCAAGTTTATAGTCAAGTGGCTCAAGTATGAA 1140
1141 GCTGCTTACATATGGAAGTCTCCACACCTGGCAACCTCCTACTGTGCTCTGGGTGAT 1200
1141 GCTGCTTACATATGGAAGTCTCCACACCTGGCAACCTCCTACTGTGCTCTGGGTGAT 1200
1201 TCTCAAGAGTCAAGCTGTGCTGGTGTCCATCTGACTTCAAAAGATTTGCTACTCT 1260
1201 TCTCAAGAGTCAAGCTGTGCTGGTGTCCATCTGACTTCAAAAGATTTGCTACTCT 1260
1261 TCTGATGACAATACACTAAAATCTGGCGCTGCAATAGAGGCTTAGAGGAGAAACAGGA 1320
1261 TCTGATGACAATACACTAAAATCTGGCGCTGCAATAGAGGCTTAGAGGAGAAACAGGA 1320

1321 GGTGATAAATCTTTCCACGGTGGGTGGCCCTCTCAGAAAGAAAAGAGTCAAGACCTGGC 1380
1321 GGTGATAAATCTTTCCACGGTGGGTGGCCCTCTCAGAAAGAAAAGAGTCAAGACCTGGC 1380
1381 CTAGTAAACAGTAACAGAGTAGCCAGAGTACTCTCTGCCAAAGCCCCCAGGGTAAAGTGAAT 1440
1381 CTAGTAAACAGTAACAGAGTAGCCAGAGTACTCTCTGCCAAAGCCCCCAGGGTAAAGTGAAT 1440
1441 CCATCCAAATTTCTCCCCCTCATCCGAGCTTGTGTGCCCCCAAGCTGTGTGGAGACTCTCCT 1500
1441 CCATCCAAATTTCTCCCCCTCATCCGAGCTTGTGTGCCCCCAAGCTGTGTGGAGACTCTCCT 1500
1501 CTTCTTTCAATATCTCTACCTGTTCTTATTAAGAACTCTCTGCCAAAGCCCCGCTCTCC 1560
1501 CTTCTTTCAATATCTCTACCTGTTCTTATTAAGAACTCTCTGCCAAAGCCCCGCTCTCC 1560
1561 ATCAACAGAGAGGCTCTGCTCTCTCTCCCAAGCCACCTTCAATCTTTCAAGATG 1620
1561 ATCAACAGAGAGGCTCTGCTCTCTCTCTCTCTCCCAAGCCACCTTCAATCTTTCAAGATG 1620
1621 TCGATTAGAAAATCTGGGTGAACCGAAACACTTCTCTCATCAACCCCATCACTCCACTGCT 1680
1621 TCGATTAGAAAATCTGGGTGAACCGAAACACTTCTCTCATCAACCCCATCACTCCACTGCT 1680
1681 TCGGAGACCAAGATCATGCTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCACTCC 1740
1681 TCGGAGACCAAGATCATGCTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCACTCC 1740
1741 CAAGCAGAGGCTTGTCTCTGAGCTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGT 1800
1741 CAAGCAGAGGCTTGTCTCTGAGCTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGT 1800
1801 CTGGAGAGTGTGAACAAAAGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGGC 1860
1801 CTGGAGAGTGTGAACAAAAGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGGC 1860
1861 CAAAGTTGAAAATCTTTTGGATCTGTGCTCTGCTTGTGTAACCCAGGAGACCTTAGT 1920
1861 CAAAGTTGAAAATCTTTTGGATCTGTGCTCTGCTTGTGTAACCCAGGAGACCTTAGT 1920
1921 AAGGACTCTTAGTCTCTACCAAAATCAAGCAAAATGGAAGAGCTGGTACCAAGTATCTCA 1980
1921 AAGGACTCTTAGTCTCTACCAAAATCAAGCAAAATGGAAGAGCTGGTACCAAGTATCTCA 1980
1981 GAGCCTCCGTCTCTCTCATCAGTCCGTATGCTTCAGAAAAGCTGTGGAAACGCTACTCTCT 2040
1981 GAGCCTCCGTCTCTCTCATCAGTCCGTATGCTTCAGAAAAGCTGTGGAAACGCTACTCTCT 2040
2041 TTGAGACCTTTGGAGAGAGGCTCTGAAATGGTGGCAAGAGAAATAGTTTCCCGAGAGAT 2100
2041 TTGAGACCTTTGGAGAGAGGCTCTGAAATGGTGGCAAGAGAAATAGTTTCCCGAGAGAT 2100
2101 AAAAACTGGTCTTGGCCATGGCAGCAAAAGGAGGCTGAGAACTCTCTCCACGAGT 2160
2101 AAAAACTGGTCTTGGCCATGGCAGCAAAAGGAGGCTGAGAACTCTCTCCACGAGT 2160
2161 CCGTCATCCAGACACCCAAATTCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGTCT 2220
2161 CCGTCATCCAGACACCCAAATTCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGTCT 2220
2221 ACCATCAGCCCGAGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCCGAGAG 2280
2221 ACCATCAGCCCGAGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCCGAGAG 2280
2281 GACTTCTGTGGTCTGAAACACTCAACAGAAATTAAGATTTCTAATCTGAGTGAAGTACTGA 2340
2281 GACTTCTGTGGTCTGAAACACTCAACAGAAATTAAGATTTCTAATCTGAGTGAAGTACTGA 2340
2341 GCTTTGGTCCACTAAAAAAGCTGAGCTTTGGTCCACTTAAACCAAGATGAAAAATACAAG 2400
2341 GCTTTGGTCCACTAAAAAAGCTGAGCTTTGGTCCACTTAAACCAAGATGAAAAATACAAG 2400

Db 1261 TCTGATGACATACATAAATCTGGCGCTTGAATAGGGCTTAGAGGAGAAACAGGA 1320
Qy 1321 GGTGATAAATTTTCCAGGTGGGTGGGCTCTCAGAGAAAAAGAGTCAAGACCTGGC 1380
Db 1321 GGTGATAAATTTTCCAGGTGGGTGGGCTCTCAGAGAAAAAGAGTCAAGACCTGGC 1380
Qy 1381 CTAGTAAACGTAAACGATGACAGAGTACTCTGCCAAAGCCCCAGGGTAAAGTGCAT 1440
Db 1381 CTAGTAAACGTAAACGATGACAGAGTACTCTGCCAAAGCCCCAGGGTAAAGTGCAT 1440
Qy 1441 CCATCCAAATTTTCCCGCTCATCCGAGCTGTGCCCCAAGCTGTGCTGGAGACCTCCCT 1500
Db 1441 CCATCCAAATTTTCCCGCTCATCCGAGCTGTGCCCCAAGCTGTGCTGGAGACCTCCCT 1500
Qy 1501 CTTCTTCAAAATACCTCTAGTTCTCTATTAATAAAGCTCTCTGCCAAGCCCGGTCTCCC 1560
Db 1501 CTTCTTCAAAATACCTCTAGTTCTCTATTAATAAAGCTCTCTGCCAAGCCCGGTCTCCC 1560
Qy 1561 ATCAACAGAGAGGCTCTGTCTCTCCGCTCTCTCCAGCCACCTTCACTTTCAAGATG 1620
Db 1561 ATCAACAGAGAGGCTCTGTCTCTCCGCTCTCTCCAGCCACCTTCACTTTCAAGATG 1620
Qy 1621 TCGATTTAGAAAATCTGGGTGACCCGAGACACCTTCTCATCACACCCATCACTCCACTGCT 1680
Db 1621 TCGATTTAGAAAATCTGGGTGACCCGAGACACCTTCTCATCACACCCATCACTCCACTGCT 1680
Qy 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCCTGTGAGCCAGAGTCACTC 1740
Db 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCCTGTGAGCCAGAGTCACTC 1740
Qy 1741 CAAGCAGAGGCTGTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGT 1800
Db 1741 CAAGCAGAGGCTGTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGT 1800
Qy 1801 CTGGAGAGTGGAACAAAGGTGTGGAAGAGTTGTAACCTGTGACTGAGCTTGATGCG 1860
Db 1801 CTGGAGAGTGGAACAAAGGTGTGGAAGAGTTGTAACCTGTGACTGAGCTTGATGCG 1860
Qy 1861 CAAGTTGAAAATCTTCAATTTGGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1861 CAAGTTGAAAATCTTCAATTTGGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Qy 1921 AAGGACTCTTAGTCTCTACCAATCAAGCAAAATTTGAAGAGCTGGTACCAAGTATCTCA 1980
Db 1921 AAGGACTCTTAGTCTCTACCAATCAAGCAAAATTTGAAGAGCTGGTACCAAGTATCTCA 1980
Qy 1981 GAGCTCTCGTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGGAACGCTACTCTTCT 2040
Db 1981 GAGCTCTCGTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGGAACGCTACTCTTCT 2040
Qy 2041 TTGAGACCTTGTGGAGAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTTCCCAGAGAT 2100
Db 2041 TTGAGACCTTGTGGAGAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTTCCCAGAGAT 2100
Qy 2101 AAAAACTGGTTGTGGCCATGGCAGCCAAAACGGAAGCTGAGAAATCCATCTCCAGAGT 2160
Db 2101 AAAAACTGGTTGTGGCCATGGCAGCCAAAACGGAAGCTGAGAAATCCATCTCCAGAGT 2160
Qy 2161 CCGTCAATCCAGACACCAATTTCCAGGACAGAGCGGAAGACATTTGCAAGCCGGTC 2220
Db 2161 CCGTCAATCCAGACACCAATTTCCAGGACAGAGCGGAAGACATTTGCAAGCCGGTC 2220
Qy 2221 ACCATCACGCCAGCTCCATGAGGAAAATCTGCACATACTTCCATAGAAAGTCCCAGGAG 2280
Db 2221 ACCATCACGCCAGCTCCATGAGGAAAATCTGCACATACTTCCATAGAAAGTCCCAGGAG 2280
Qy 2281 GACTTCTGTGCTGAAACACTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTACTGA 2340
Db 2281 GACTTCTGTGCTGAAACACTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTACTGA 2340
Qy 2341 GCTTTGGTCCACTTAAACAGCTGAGCTTTGGTCCACTAAACAGATGAAAAATACAG 2400

Db 2341 GCTTTGGTCCACTTAAACAGCTGAGCTTTGGTCCACTTAAACAGATGAAAAATACAG 2400
Qy 2401 AGTGACTCTATAACTCTGGTCTTTAAGAAAGCTGCTTTTCATTTTATAGAAAAATCTTT 2460
Db 2401 AGTGACTCTATAACTCTGGTCTTTAAGAAAGCTGCTTTTCATTTTATAGAAAAATCTTT 2460
Qy 2461 TCAACGCTGGAATGTACCTAATCTGGTCTTACTACTACCAATAATGTATATGCAGCTTCCCAG 2520
Db 2461 TCAACGCTGGAATGTACCTAATCTGGTCTTACTACTACCAATAATGTATATGCAGCTTCCCAG 2520
Qy 2521 GATGAATGCTGTGTTTAAATTTCAAAAGTAAATTTGTCACTCTAGCAATTTTGAATGAAT 2580
Db 2521 GATGAATGCTGTGTTTAAATTTCAAAAGTAAATTTGTCACTCTAGCAATTTTGAATGAAT 2580
Qy 2581 AGTCTTCACTTTTAAATTTTCAATCTCTCTATAATAATGACATCCAGTTCATGGAGG 2640
Db 2581 AGTCTTCACTTTTAAATTTTCAATCTCTCTATAATAATGACATCCAGTTCATGGAGG 2640
Qy 2641 CAAAAACAAAGTTTCTGTTATCTGAAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAG 2700
Db 2641 CAAAAACAAAGTTTCTGTTATCTGAAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAG 2700
Qy 2701 CCACAGCATGAGGCTGTGAAAGGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCT 2760
Db 2701 CCACAGCATGAGGCTGTGAAAGGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCT 2760
Qy 2761 GTTCTGCTCCAAACATGTATAATTTTATTGAAATACATAATCTTTTCACTATGAAAAA 2820
Db 2761 GTTCTGCTCCAAACATGTATAATTTTATTGAAATACATAATCTTTTCACTATGAAAAA 2820
Qy 2821 AAAAAAATAA 2831
Db 2821 AAAAAAATAA 2831

RESULT 3
US-09-780-053-1
; Sequence 1, Application US/097800053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E. H. Afar
; APPLICANT: Pia M. Challita-Bid
; APPLICANT: Mary Paris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; FILE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 129,5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2838
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (130)...(2322)
US-09-780-053-1

Query Match 99.5%; Score 2816.2; DB 3; Length 2838;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2821; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 3 CACGAGCGGGAGTTGGAGCGGCAATACGATTTGTGTTGAGAGCGGCAACGTGCCATTTTC 62
Db 9 CAGTGGCGGGAGTTGGAGCGGCAATACGATTTGTGTTGAGAGCGGCAACGTGCCATTTTC 68

QY 63 TGCTGAACCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCT 122
Db 69 TGCTGAACCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCT 128
QY 123 GATGCTCTTCAATTCGGTGTCTCGCCAGGCCAGCTTGGCGTCTCAGAAATGGATGGTC 182
Db 129 GATGCTCTTCAATTCGGCGCTCGCCAGGCCAGCTTGGCGTCTCAGAAATGGATGGTC 188
QY 183 TTCACAATACCTCTTCAATCCCTCTGACTGGTTATCAGTCAGTGGTAAATGATGAACA 242
Db 189 TTCACAATACCTCTTCAATCCCTCTGACTGGTTATCAGTCAGTGGTAAATGATGAACA 248
QY 243 CACTTCTTATGGAGAACAGGAGTCCAGTCTCTCTTTTGGATGATCCTTCTCTG 302
Db 249 CACTTCTTATGGAGAACAGGAGTCCAGTCTCTCTTTTGGATGATCCTTCTCTG 308
QY 303 TCCCAATATGGAAATGATGATAGCAGTTGCAATGAAGAAGCTTTTGGTTCGATGATATA 362
Db 309 TCCCAATATGGAAATGATGATAGCAGTTGCAATGAAGAAGCTTTTGGTTCGATGATATA 368
QY 363 CACAGAATCACAAAGTTTCAGAAAGAGTCTTCAAGATGGATGGCTCACTGGATGC 422
Db 369 CACAGAATCACAAAGTTTCAGAAAGAGTCTTCAAGATGGATGGCTCACTGGATGC 428
QY 423 CGTCTTTGACCTGGCTGGGTTCTCGGTGAACCTTAACTTTGTTACAGCAGCAGGTGATCA 482
Db 429 CGTCTTTGACCTGGCTGGGTTCTCGGTGAACCTTAACTTTGTTACAGCAGCAGGTGATCA 488
QY 483 AACAGCCAAATTTTGGGAGCTAAAGCTGGTGAAGTGAATGGAAACATGCAAAAGGTGATCA 542
Db 489 AACAGCCAAATTTTGGGAGCTAAAGCTGGTGAAGTGAATGGAAACATGCAAAAGGTGATCA 548
QY 543 ATGCAGCTCAAGTCAAGTGGCTTTTCTAAAGTTGAGAAAGCTGATTTCTGACGGGTG 602
Db 549 ATGCAGCTCAAGTCAAGTGGCTTTTCTAAAGTTGAGAAAGCTGATTTCTGACGGGTG 608
QY 603 AAGAGATGGCAACATTTATGCTCTGGGATACAGGTGCAACAAAGAGATGGGTTTATAG 662
Db 609 AAGAGATGGCAACATTTATGCTCTGGGATACAGGTGCAACAAAGAGATGGGTTTATAG 668
QY 663 GCAAGTGAATCAATCAGTGGAGCTCAATACCTCAGACAAGCAAAACCCCTTCAAAACC 722
Db 669 GCAAGTGAATCAATCAGTGGAGCTCAATACCTCAGACAAGCAAAACCCCTTCAAAACC 728
QY 723 CAAGAAGAAACAGAAATCAAAAGGACTTGTCTCTCTGTTGATTTCCAGCAAGGTGATAC 782
Db 729 CAAGAAGAAACAGAAATCAAAAGGACTTGTCTCTCTGTTGATTTCCAGCAAGGTGATAC 788
QY 783 TGTGGTCTCTTTCAAGAAGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAAT 842
Db 789 TGTGGTCTCTTTCAAGAAGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAAT 848
QY 843 CAAAGATATGGGATTTACGTAAGAAATTTATCTGCTTATCGACAAGAACCCATAGCATCCAA 902
Db 849 CAAAGATATGGGATTTACGTAAGAAATTTATCTGCTTATCGACAAGAACCCATAGCATCCAA 908
QY 903 GTCTTTCTGTACCCAGGTAGAGCACTCGAAGAACTTGGATATTCAGTCTGATTTTGA 962
Db 909 GTCTTTCTGTACCCAGGTAGAGCACTCGAAGAACTTGGATATTCAGTCTGATTTTGA 968
QY 963 TTCCACTGGCTCTACTTTTATTTGCTAAATGTCACAGACGATTAACATCTACATGTTTAAAT 1022
Db 969 TTCCACTGGCTCTACTTTTATTTGCTAAATGTCACAGACGATTAACATCTACATGTTTAAAT 1028
QY 1023 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTA 1082
Db 1029 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTA 1088
QY 1083 TGTAAATCCAGCTTAGTCCAGATGACCAAGTTTATGTCAGTGGCTCAAGTATGAAGC 1142
Db 1089 TGTAAATCCAGCTTAGTCCAGATGACCAAGTTTATGTCAGTGGCTCAAGTATGAAGC 1148
QY 1143 TGCTACATATGGAAGGTCTCCACACCTGGCAACCTCTCTACTGTGCTCTCTGGTCAATTC 1202

Db 1149 TGCTACATATGGAAGGTCTCCACACCTTGGCAACCTCTCTACTGTGCTCTCTGGTCAATTC 1208
QY 1203 TCAAGAGGTCAAGTCTGTGTGCTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1262
Db 1209 TCAAGAGGTCAAGTCTGTGTGCTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1268
QY 1263 TGATGACAAATACATCAATAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGAGG 1322
Db 1269 TGATGACAAATACATCAATAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGAGG 1328
QY 1323 TGATAAACTTTCCACGGTGGTGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACCTTGGGCT 1382
Db 1329 TGATAAACTTTCCACGGTGGTGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACCTTGGGCT 1388
QY 1383 AGTAACAGTAAGAGTAGCCAGAGTACTCTCTGCCAAAGCCCGGTAAGAGTGAATCC 1442
Db 1389 AGTAACAGTAAGAGTAGCCAGAGTACTCTCTGCCAAAGCCCGGTAAGAGTGAATCC 1448
QY 1443 ATCCAAATCTTCCCGCTCATCCGAGCTTGTGCCCAAGCTGTCTGGAGAGCTTCCCTCT 1502
Db 1449 ATCCAAATCTTCCCGCTCATCCGAGCTTGTGCCCAAGCTGTCTGGAGAGCTTCCCTCT 1508
QY 1503 TCCCTTCAATACCTCTACGTTCTCTATTAAAACTCTCTGCAAGGCCCGGTCTCCCAT 1562
Db 1509 TCCCTTCAATACCTCTACGTTCTCTATTAAAACTCTCTGCAAGGCCCGGTCTCCCAT 1568
QY 1563 CAAAGAGAGGCTCTGTCTCTCTCCAGCAACCTTCTCTCAAGCCACCTTCTCTTTCAAGATGC 1622
Db 1569 CAAAGAGAGGCTCTGTCTCTCTCCAGCAACCTTCTCTCAAGCCACCTTCTCTTTCAAGATGC 1628
QY 1623 GATTAGAACTGGGTGACCCGAAACCTTCTCTCATCACCCACCTCACTCCACCTGCTTC 1682
Db 1629 GATTAGAACTGGGTGACCCGAAACCTTCTCTCATCACCCACCTCACTCCACCTGCTTC 1688
QY 1683 GGAGACCAGATCATGTCTCCGAGAAAGCCCTTATCTCTGTGAGCCAGAAAGTCAATCCA 1742
Db 1689 GGAGACCAGATCATGTCTCCGAGAAAGCCCTTATCTCTGTGAGCCAGAAAGTCAATCCA 1748
QY 1743 AGCAGAGGCTTGTCTGTAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT 1802
Db 1749 AGCAGAGGCTTGTCTGTAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT 1808
QY 1803 GGAGAGTGTGAACAAAGTGTGCAAGATTTAACTGTGTGACTGTGAGCTGTGATGGCCA 1862
Db 1809 GGAGAGTGTGAACAAAGTGTGCAAGATTTAACTGTGTGACTGTGAGCTGTGATGGCCA 1868
QY 1863 AGTTGAAAATCTTCAATTTGGATCTGTGCTGCTTCTGTGTAAACAGGAAGACCTTAGTAA 1922
Db 1869 AGTTGAAAATCTTCAATTTGGATCTGTGCTGCTTCTGTGTAAACAGGAAGACCTTAGTAA 1928
QY 1923 GGACTCTCTAGGCTCTCAATCAAGCAAAATGGAAGAGCTGGTACAGATATCTCAGA 1982
Db 1929 GGACTCTCTAGGCTCTCAATCAAGCAAAATGGAAGAGCTGGTACAGATATCTCAGA 1988
QY 1983 GCCTCGCTCTCTATCAGTCCGTTATGCTTCAAGAGCTGTGGAACGCTCTCTCTTT 2042
Db 1989 GCCTCGCTCTCTATCAGTCCGTTATGCTTCAAGAGCTGTGGAACGCTCTCTCTTT 2048
QY 2043 GAGACTTGTGGAAGGCTCTGAAATGGTAGCAAGAGAAATAGTTTCCAGAGAAATAA 2102
Db 2049 GAGACTTGTGGAAGGCTCTGAAATGGTAGCAAGAGAAATAGTTTCCAGAGAAATAA 2108
QY 2103 AAACTGGTGTGGCCATGGCAGCAACCGAAGGCTGAGATTCATCTCCAGAGAGTCC 2162
Db 2109 AAACTGGTGTGGCCATGGCAGCAACCGAAGGCTGAGATTCATCTCCAGAGAGTCC 2168
QY 2163 GTCATCCAGACACCAATTTCCAGAGAGCAGAGCGGAAGACATTTGCCAAGCCCGGTAC 2222
Db 2169 GTCATCCAGACACCAATTTCCAGAGAGCAGAGCGGAAGACATTTGCCAAGCCCGGTAC 2228
QY 2223 CATCAGCCAGCTCCATGAGGAAATCTGACATATCTTCCATAGAAAAGTCCAGAGAGA 2282

```
Db 2229 CATCAGCCCGCTCCATGAGGAAATCTGCACATACCTTCCATAGAAAGTCCCGAGGA 2288
Qy 2283 CTTCTGTGTCTGAACACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGC 2342
Db 2289 CTTCTGTGTCTGAACACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGC 2348
Qy 2343 TTTGTGCCATAAACAAGCTGAGCTTGGTCCACTTAAACAGATGAAATATACAGAG 2402
Db 2349 TTTGTGCCATAAACAAGCTGAGCTTGGTCCACTTAAACAGATGAAATATACAGAG 2408
Qy 2403 TGACTCTATAACTCTGGTCTTTAAGAAAGCTGCTTTTCATTTTATAGACAAATCTTTTC 2462
Db 2409 TGACTCTATACTCTGGTCTTTAAGAAAGCTGCTTTTCATTTTATAGACAAATCTTTTC 2468
Qy 2463 AACGCTGAAATGTAACCTAATCTGGTCTTACTACTACCAATATGATATGCGAGCTTCCCGAGGA 2522
Db 2469 AACGCTGAAATGTAACCTAATCTGGTCTTACTACTACCAATATGATATGCGAGCTTCCCGAGGA 2528
Qy 2523 TGAATGCTGTGTTAAATTTCAATAGTAAATTTGTCACTCTAGCAATTTTGAATGAATAG 2582
Db 2529 TGAATGCTGTGTTAAATTTCAATAGTAAATTTGTCACTCTAGCAATTTTGAATGAATAG 2588
Qy 2583 TCTTCACTTTTAAATTTATCTTCTCTATAATAATGACATCCAGTTTATGAGGCA 2642
Db 2589 TCTTCACTTTTAAATTTATCTTCTATAATAATGACATCCAGTTTATGAGGCA 2648
Qy 2643 AAAAAACAAGTTCTTGTGTTATCTCGAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702
Db 2649 AAAAAACAAGTTCTTGTGTTATCTCGAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2708
Qy 2703 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCTCTGCTGAAGACCCCTGTTCTGT 2762
Db 2709 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCTCTGCTGAAGACCCCTGTTCTGT 2768
Qy 2763 TCTGCTCCAAATGATATATTTTATTTGAAATACATAATCTTTTCACTATGAAAAAAA 2822
Db 2769 TCTGCTCCAAATGATATATTTTATTTGAAATACATAATCTTTTCACTATGAAAAAAA 2828
Qy 2823 AAAAAAAA 2831
Db 2829 AAAAAAAA 2837
```

```
RESULT 4
US-09-814-353-20186
; Sequence 20186, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20186
; LENGTH: 4422
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 4406, 4407, 4408, 4409, 4410, 4411, 4412, 4413, 4414, 4415,
; LOCATION: 4416, 4417, 4418, 4419, 4420, 4421, 4422
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20186

Query Match 99.0%; Score 2804; DB 3; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CACGAGCGGAGTTGGAGCGCATACGATTTTGTGTGTGAGAGGGCGAACCTGGGATTTTC 62
Db 8 CAGTGGCGGAGTTGGAGCGCATACGATTTTGTGTGTGAGAGGGCGAACCTGGGATTTTC 67
Qy 63 TGTGNACTTCGAGGCACTTTCTAGACATTTTCTCTCAGCTCAGGCTTTTCTCCGACCT 122
Db 68 TGTGNACTTCGAGGCACTTTCTAGACATTTTCTCAGCTCAGGCTTTTCTCCGACCT 127
Qy 123 GATGCTCTTCAATTCGGTGTCTCCGACCCAGCGCTTGGCGTCTTGAGAAATGGAATGTC 182
Db 128 GATGCTCTTCAATTCGGTGTCTCCGACCCAGCGCTTGGCGTCTTGAGAAATGGAATGTC 187
Qy 193 TTCAATACCTCTTCAATCCCTTCTGACTTGTGATTCAGTGCAGTGGTATGATGAACA 242
Db 188 TTCAATACCTCTTCAATCCCTTCTGACTTGTGATTCAGTGCAGTGGTATGATGAACA 247
Qy 243 CACTTCTTATGAGAAACAGAGTCCAGTTCCTCTCTTTTGGATGTACCTTCTCTTCGC 302
Db 248 CACTTCTTATGAGAAACAGAGTCCAGTTCCTCTCTTTTGGATGTACCTTCTCTTCGC 307
Qy 303 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAGCTTTGTTCGATTTATAA 362
Db 308 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAGCTTTGTTCGATTTATAA 367
Qy 363 CACAGATCAAAAATTTTCAGAAAGAGTGTCTCAAAAGATGGTCTCACTGCGAATGC 422
Db 368 CACAGATCAAAAATTTTCAGAAAGAGTGTCTCAAAAGATGGTCTCACTGCGAATGC 427
Qy 423 CGTCTTTGACCTGGCTGGGTTCTGTTGTAACCTTTAACTTTTACAGCAGCAGGTGATCA 482
Db 428 CGTCTTTGACCTGGCTGGGTTCTGTTGTAACCTTTAACTTTTACAGCAGCAGGTGATCA 487
Qy 483 AACAGCCAAATTTTGGGACGTAAAGCTGGTGAATGGAACATGCAAGGTCATCA 542
Db 488 AACAGCCAAATTTTGGGACGTAAAGCTGGTGAATGGAACATGCAAGGTCATCA 547
Qy 543 ATGACGCTCAAGTCAGTTGCTTTTCTAAGTTTGAAGAGCTGTATTTCTGACGGTGG 602
Db 548 ATGACGCTCAAGTCAGTTGCTTTTCTAAGTTTGAAGAGCTGTATTTCTGACGGTGG 607
Qy 603 AAGAGATGGCAACATTTATGGTCTGGGATACAGGTCACAAAGAGATGGGTTTATAG 662
Db 608 AAGAGATGGCAACATTTATGGTCTGGGATACAGGTCACAAAGAGATGGGTTTATAG 667
Qy 663 GCAAGTGAATCAAAATCAGTGGAGCTCACAATACCTCAGCAAGCAAAACCCCTTCAAAACC 722
Db 668 GCAAGTGAATCAAAATCAGTGGAGCTCACAATACCTCAGCAAGCAAAACCCCTTCAAAACC 727
Qy 723 CAAGAAAGAAACAGAAATTCAAAAGAGCTTCTCTCTGTTGTTTCCAGCAAGGTATAC 782
Db 728 CAAGAAAGAAACAGAAATTCAAAAGAGCTTCTCTCTGTTGTTTCCAGCAAGGTATAC 787
Qy 783 TGTGTCCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAT 842
Db 788 TGTGTCCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAT 847
Qy 843 CAAAGTATGGGATTTACGTAAGAAATTTATCTGCTTTATCGAAGAAACCCATAGCATCCA 902
Db 848 CAAAGTATGGGATTTACGTAAGAAATTTATCTGCTTTATCGAAGAAACCCATAGCATCCA 907
```


Qy	903	GTCTTTTCTCTGTATCCAGAGGTAGCAGCACTCGAAAACTTTGGATATTTCAAGTCTGATTTTGGAA	962
Db	908	GTCTTTTCTCTGTATCCAGAGGTAGCAGCACTCGAAAACTTTGGATATTTCAAGTCTGATTTTGGAA	967
Qy	963	TTCCACATCGGCTCTACTTTTATTTTGTCTAATTCACACAGACGATAAACAATCTACATGTTTAAATAT	1022
Db	968	TTCCACATCGGCTCTACTTTTATTTTGTCTAATTTGCAACAGACGATAAACAATCTACATGTTTAAATAT	1027
Qy	1023	GACTGGGTTGAAGACTTCTCCAGTGGGCTATTTTCCAAATGGGACACAGAACTCTACCTTTTAA	1082
Db	1028	GACTGGGTTGAAGACTTCTCCAGTGGGCTATTTTCCAAATGGGACACAGAACTCTACCTTTTAA	1087
Qy	1083	TGTAATAATCCAGCCCTTAGTCAGATGACAGTGTTTTATAGTCAGTGGCTCAAGTGAATGAAGC	1142
Db	1088	TGTAATAATCCAGCCCTTAGTCAGATGACAGTGTTTTATAGTCAGTGGCTCAAGTGAATGAAGC	1147
Qy	1143	TGCCCTACATATGGAAGGTCCTCCACACCCCTGGCAACCTCTTACTGTGCTCTCGTGGGTCAATTC	1202
Db	1148	TGCCCTACATATGGAAGGTCCTCCACACCCCTGGCAACCTCTTACTGTGCTCTCGTGGGTCAATTC	1207
Qy	1203	TCAGAGGTCACGTCCTGTGTGCTGGTGTCCATCTGACTTTCACAAAGATTTGCTACCTGTGTC	1262
Db	1208	TCAGAGGTCACGTCCTGTGTGCTGGTGTCCATCTGACTTTCACAAAGATTTGCTACCTGTGTC	1267
Qy	1263	TGATGACAATACACTAAATAATCTGGCCCTTGAAATAGAGGCTTAGAGAGAGAAAACGAGGAGG	1322
Db	1268	TGATGACAATACACTAAATAATCTGGCCCTTGAAATAGAGGCTTAGAGAGAGAAAACGAGGAGG	1327
Qy	1323	TGATAAACTTTCCACGGTGGGTTGGGGCTCTCGAGAGAAAAGAGTCAAGACCTGGCCCT	1382
Db	1328	TGATAAACTTTCCACGGTGGGTTGGGGCTCTCGAGAGAAAAGAGTCAAGACCTGGCCCT	1387
Qy	1383	AGTAAACAGTAAACAGTAGTCCAGAGTACTCTCTGCCAAAGCCCCACAGGTAAGTGCATATCC	1442
Db	1388	AGTAAACAGTAAACAGTAGTCCAGAGTACTCTCTGCCAAAGCCCCACAGGTAAGTGCATATCC	1447
Qy	1443	ATCCAAATCTTCCCGTCTATCCGAGCTTGTGGCCCAAGCTGTGTGCGAGACCTCTCCCTCT	1502
Db	1448	ATCCAAATCTTCCCGTCTATCCGAGCTTGTGGCCCAAGCTGTGTGCGAGACCTCTCCCTCT	1507
Qy	1503	TCCTTCAAAATACCTCTACGTTCTCTATTAATAAACCCTCTCTGCCAAGGCCCGGTCTCCCAT	1562
Db	1508	TCCTTCAAAATACCTCTACGTTCTCTATTAATAAACCCTCTCTGCCAAGGCCCGGTCTCCCAT	1567
Qy	1563	CAACAGAAAGAGGCTCTGTCTCTCCGTCCTCTCCAAAGCCACCTTCACTCTTTCAAGATGTC	1622
Db	1568	CAACAGAAAGAGGCTCTGTCTCTCCGTCCTCTCCAAAGCCACCTTCACTCTTTCAAGATGTC	1627
Qy	1623	GATTAGAAAATGGGTGACCCGAAACAACCTTCTCATCAACCCCATCACTCCACTCTGCTTC	1682
Db	1628	GATTAGAAAATGGGTGACCCGAAACAACCTTCTCATCAACCCCATCACTCCACTCTGCTTC	1687
Qy	1683	GGAGACCAAGATCATGTCTCCGAGAAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCA	1742
Db	1688	GGAGACCAAGATCATGTCTCCGAGAAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCA	1747
Qy	1743	AGCAGAGGCTTGCTCTCAGTCTAGAAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT	1802
Db	1748	AGCAGAGGCTTGCTCTCAGTCTAGAAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT	1807
Qy	1803	GGAGAGTGTGAAAACAAAAGTGTGTGAAGAGTGTGTAACTGTGTGACTTGAGCTTGAATGGCCA	1862
Db	1808	GGAGAGTGTGAAAACAAAAGTGTGTGAAGAGTGTGTAACTGTGTGACTTGAGCTTGAATGGCCA	1867
Qy	1863	AGTTGAAAATCTTCAATTTGGATCTGTGCTCTGCCCTGTGGTAAACCGAGGAAGACTTAGTAA	1922
Db	1868	AGTTGAAAATCTTCAATTTGGATCTGTGCTCTGCCCTGTGGTAAACCGAGGAAGACTTAGTAA	1927
Qy	1923	GGACTCTCTAGGTCCTACCAAAATCAAGCAAAAATTTGAAGGAGCTGGTACCAGTATCTCAGA	1982
Db	1928	GGACTCTCTAGGTCCTACCAAAATCAAGCAAAAATTTGAAGGAGCTGGTACCAGTATCTCAGA	1987
Qy	1983	GCCTCCGTCCTTATCAGTCCGATGTGCTTCAGAAAAGCTGTGGAAACGCTACTCTCTCCCTTT	2042

Db	1988	GCCTCCGCTCCTATCAGTCGGTATGCTTCAGAAAGCTGTGGAAACGCTACTCTTCTCTTT	2047
Qy	2043	GAGACCTTGTGGAGAAAGGCTCGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAATAA	2102
Db	2048	GAGACCTTGTGGAGAAAGGCTCGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAATAA	2107
Qy	2103	AAACTGGTGTGTGGCCATGGCAGCAACAGGAAGGCTGAGAAATCCATCTCCACGAAGTCC	2162
Db	2108	AAACTGGTGTGTGGCCATGGCAGCAACAGGAAGGCTGAGAAATCCATCTCCACGAAGTCC	2167
Qy	2163	GTCATCCAGACACCCCAATTTCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTTCAC	2222
Db	2168	GTCATCCAGACACCCCAATTTCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTTCAC	2227
Qy	2223	CATCACGCCAGGCTCCATGAGGAAATCTCCACATATCTCCATAGAAGTCCCAAGGAGGA	2282
Db	2228	CATCACGCCAGGCTCCATGAGGAAATCTCCACATATCTCCATAGAAGTCCCAAGGAGGA	2287
Qy	2283	CTTCTGTGGTCTGAAACACTCAACAGAGAAATATAGATTTCTAATCTGTAGTGAGTTACTGAGC	2342
Db	2288	CTTCTGTGGTCTGAAACACTCAACAGAGAAATATAGATTTCTAATCTGTAGTGAGTTACTGAGC	2347
Qy	2343	TTTGGTCCACTAAACCAAGCTGAGCTTTGGTCCACTAAACCAAGATGAAATAACAAGAG	2402
Db	2348	TTTGGTCCACTAAACCAAGCTGAGCTTTGGTCCACTAAACCAAGATGAAATAACAAGAG	2407
Qy	2403	TGACTCTATAACTCTGTCTTTAAGAAAGCTGCCCTTTTCATTTTACACAAATCTTTTC	2462
Db	2408	TGACTCTATAACTCTGTCTTTAAGAAAGCTGCCCTTTTCATTTTACACAAATCTTTTC	2467
Qy	2463	AACGCTGAAATGTAACCTTAATCTGGTCTCTACTACCATAATGTATATGCAGCTTCCCGAGGA	2522
Db	2468	AACGCTGAAATGTAACCTTAATCTGGTCTCTACTACCATAATGTATATGCAGCTTCCCGAGGA	2527
Qy	2523	TGAATGCTGTGTTTAAATTTTATAAAGTAAATTTGTCACTCTAGCAATTTTGAATGAATAG	2582
Db	2528	TGAATGCTGTGTTTAAATTTTATAAAGTAAATTTGTCACTCTAGCAATTTTGAATGAATAG	2587
Qy	2583	TCCTTCACTTTTAAATTAATTCATCTCTCTATAATAATGACATCCCAAGTTTCATGGAGGCA	2642
Db	2588	TCCTTCACTTTTAAATTAATTCATCTCTCTATAATAATGACATCCCAAGTTTCATGGAGGCA	2647
Qy	2643	AAAAACAAGTTTCTGTATATCCCTGAAACTTTCTATGCTCAGTGAAAGTATCTGCCAGCC	2702
Db	2648	AAAAACAAGTTTCTGTATATCCCTGAAACTTTCTATGCTCAGTGAAAGTATCTGCCAGCC	2707
Qy	2703	ACAGCATGAGGCCCTGTGAAGGCTGACGTGAGAAATCTCTGCTGAAGACCCCTGGTTCCTGT	2762
Db	2708	ACAGCATGAGGCCCTGTGAAGGCTGACGTGAGAAATCTCTGCTGAAGACCCCTGGTTCCTGT	2767
Qy	2763	TCCTGCCCTCCACATGTATAATTTTAAATTTGAAATAACATAAATCTTTTCTACTATG	2814
Db	2768	TCCTGCCCTCCACATGTATAATTTTAAATTTGAAATAACATAAATCTTTTCTACTATG	2819

RESULT 5

```

US-10-357-930-25699
; Sequence 25699, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSIT
; TITLE OF INVENTION: IDENTIFICATION ASSE
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319

```

```
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 6232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25699
; LENGTH: 4422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4406, 4407, 4408, 4409, 4410, 4411, 4412, 4413, 4414, 4415,
; LOCATION: 4416, 4417, 4418, 4419, 4420, 4421, 4422
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-25699

Query Match          99.0%; Score 2804; DB 9; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACGAGCGGGAGTTGGAGCGGATAACGATTTGTGTGTGAGAGGGGCAACGTGCGGATTTTC 62
   |||
Db 8 CAGTGGCGGGAGTTGGAGCGGATAACGATTTGTGTGTGAGAGGGGCAACGTGCGGATTTTC 67

QY 63 TGTGTAACCTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTTCGACACCTT 122
   |||
Db 68 TGTGTAACCTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTTCGACACCTT 127

QY 123 GATGCTCTTCAATTCGGTGTCTCGGCCAGCCAGCTTGGCGTCTCGAGAAATGGATGGTC 182
   |||
Db 128 GATGCTCTTCAATTCGGTGTCTCGGCCAGCCAGCTTGGCGTCTCGAGAAATGGATGGTC 187

QY 183 TTCACAAATACCTCTTCAATCCCTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 242
   |||
Db 188 TTCACAAATACCTCTTCAATCCCTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 247

QY 243 CACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTTGGATGATCTTCTCTCTGTC 302
   |||
Db 248 CACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTTGGATGATCTTCTCTCTGTC 307

QY 303 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGGAAGGCTTTGTTTCGATTTGTATAA 362
   |||
Db 308 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGGAAGGCTTTGTTTCGATTTGTATAA 367

QY 363 CACAGAAATCAAAAGTTTCAGAAAGAAAGTCTTCAAAGAAATGGATGGCTCACTGGAAATGC 422
   |||
Db 368 CACAGAAATCAAAAGTTTCAGAAAGAAAGTCTTCAAAGAAATGGATGGCTCACTGGAAATGC 427

QY 423 CGTCTTTGACCTGGCTGGGTTCTCGGTGAACTTAAACTTTGTATACAGCAGCAGGTGATCA 482
   |||
Db 428 CGTCTTTGACCTGGCTGGGTTCTCGGTGAACTTAAACTTTGTATACAGCAGCAGGTGATCA 487

QY 483 AACAGCCAAATTTGGGACGTAAAGCTGGTGAAGCTGATTTGGAAATGCAAAAGGTCATCA 542
   |||
Db 488 AACAGCCAAATTTGGGACGTAAAGCTGGTGAAGCTGATTTGGAAATGCAAAAGGTCATCA 547

QY 543 ATGCAAGCTCAAGTCAGTTGGCTTTTCTAAGTTTGAAGAAAGCTGTATTTCTGTACGGGTGG 602
   |||
Db 548 ATGCAAGCTCAAGTCAGTTGGCTTTTCTAAGTTTGAAGAAAGCTGTATTTCTGTACGGGTGG 607

QY 603 AAGAGATGGCAACATTTATGTTCTGGATACCAAGGTGCAACAAAGAAAGATGGGTTTATAG 662
   |||
Db 608 AAGAGATGGCAACATTTATGTTCTGGATACCAAGGTGCAACAAAGAAAGATGGGTTTATAG 667

QY 663 GCNAAGTGAATCAATCAGTGGAGCTCACAAATACCTCAGACAGCAACCCCTTCAAACACC 722
   |||
```

```
Db 668 GCAAGTGAATCAATCAGTGGAGCTCACAAATACCTCAGACAAAGCAAAACCCCTTCAAACACC 727
   |||
QY 723 CAAGAAGAAACAGAAATTCAAAGAGCTTGTCTCTTGTGTGATTTTCAGGAAAGTGTTC 782
   |||
Db 728 CAAGAAGAAACAGAAATTCAAAGAGCTTGTCTCTTGTGTGATTTTCAGGAAAGTGTTC 787
   |||
QY 783 TGTGGTCTCTTTCAAGACGAGAAATACCTTGTAGTCTCAGCAGGAGCTGTGGATGGGATAT 842
   |||
Db 788 TGTGGTCTCTTTCAAGACGAGAAATACCTTGTAGTCTCAGCAGGAGCTGTGGATGGGATAT 847
   |||
QY 843 CAAGATATGGGATTTACGTAAGAAATATATCTGCTTATCGACAAGAACCCATAGCATCCAA 902
   |||
Db 848 CAAGATATGGGATTTACGTAAGAAATATATCTGCTTATCGACAAGAACCCATAGCATCCAA 907
   |||
QY 903 GTCTTTCTGTACCCAGGTAGCAGACTTCGAAAACTTGGATATTCAGACTGTGATTTTGGGA 962
   |||
Db 908 GTCTTTCTGTACCCAGGTAGCAGACTTCGAAAACTTGGATATTCAGACTGTGATTTTGGGA 967
   |||
QY 963 TTCCACTGGCTCTACTTTTATTTGCTTAATTCACAGACGATTAACATCTACATGTTTAAATAT 1022
   |||
Db 968 TTCCACTGGCTCTACTTTTATTTGCTTAATTCACAGACGATTAACATCTACATGTTTAAATAT 1027
   |||
QY 1023 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCCAGAACTCTACCTTTTA 1082
   |||
Db 1028 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCCAGAACTCTACCTTTTA 1087
   |||
QY 1083 TGTAAATCCAGCCTTAGTCCAGATGACCAAGTCTTTTGTAGTGTGCTCAAGTGTGAAGC 1142
   |||
Db 1088 TGTAAATCCAGCCTTAGTCCAGATGACCAAGTCTTTTGTAGTGTGCTCAAGTGTGAAGC 1147
   |||
QY 1143 TGCCTACATATGGAAGGTCTCCACACCTTGGCAACCTCTACTGTCTCTGGGTCAATTC 1202
   |||
Db 1148 TGCCTACATATGGAAGGTCTCCACACCTTGGCAACCTCTACTGTCTCTGGGTCAATTC 1207
   |||
QY 1203 TCAAGAGGTCACTGTGTGTCTGTGTCTCACTGTCTCACTGTCTCACTGTCTCACTGTCT 1262
   |||
Db 1208 TCAAGAGGTCACTGTGTGTCTGTGTCTCACTGTCTCACTGTCTCACTGTCTCACTGTCT 1267
   |||
QY 1263 TGATGACAAATACACTTAAATAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGAGG 1322
   |||
Db 1268 TGATGACAAATACACTTAAATAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGAGG 1327
   |||
QY 1323 TGATAACTTTTCCAGGTGGGCTTGGGCTCTCAGAGAAAGAAAGTCAAGACCTGGGCT 1382
   |||
Db 1328 TGATAACTTTTCCAGGTGGGCTTGGGCTCTCAGAGAAAGAAAGTCAAGACCTGGGCT 1387
   |||
QY 1383 AGTAACAGTAAACAGTAGCAGAGATCTCTCTGCTCCAAAGCCGCCAGGGTAAAGTGCAATCC 1442
   |||
Db 1388 AGTAACAGTAAACAGTAGCAGAGATCTCTCTGCTCCAAAGCCGCCAGGGTAAAGTGCAATCC 1447
   |||
QY 1443 ATCCAAATTTTCCCGCTCATCCGAGCTTGTGTGCTCCCAAGCTGTGTGAGAGCTTCCCTCT 1502
   |||
Db 1448 ATCCAAATTTTCCCGCTCATCCGAGCTTGTGTGCTCCCAAGCTGTGTGAGAGCTTCCCTCT 1507
   |||
QY 1503 TCCTTCAATACTCTAGCTTCTCTATTAATAACCTCTCTCTGCTCCAAAGCCGCCAGGGTCCCAT 1562
   |||
Db 1508 TCCTTCAATACTCTAGCTTCTCTATTAATAACCTCTCTCTGCTCCAAAGCCGCCAGGGTCCCAT 1567
   |||
QY 1563 CAACAAGAGAGCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1622
   |||
Db 1568 CAACAAGAGAGCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1627
   |||
QY 1623 GATTAGAAATCTGGGTGACCCGAAACACTTCTCTATCATCACCAATCTCTCTCTCTCTCTCTCTCT 1682
   |||
Db 1628 GATTAGAAATCTGGGTGACCCGAAACACTTCTCTATCATCACCAATCTCTCTCTCTCTCTCTCTCT 1687
   |||
QY 1683 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTCTGTGAGCCAGAGATCATCCCA 1742
   |||
Db 1688 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTCTGTGAGCCAGAGATCATCCCA 1747
   |||
QY 1743 AGCAGAGGCTTGTCTGTAGTCTTAGAAATAGAGTAAGAGGAGGCTTAGACTCAAGCTGTCT 1802
   |||
```


Db 721 AAAGNACTTGGCTTCTCTGTGGATTCCAGAAAGTGTACTGTGTCTCTTTCAAGAC 780
Qy 802 GAGATACCTTAGTCTCAGCAGAGCTGTGGATGGGATTAATCAAAATATGGGATTTACGT 861
Db 781 GAGATACCTTAGTCTCAGCAGAGCTGTGGATGGGATTAATCAAAATATGGGATTTACGT 840
Qy 862 AAGATTATACCTTAGTCTCAGCAGAGCCATAGCATCCAGTCTTCTCTGTACCCAGGT 921
Db 841 AAGATTATACCTTAGTCTCAGCAGAGCCATAGCATCCAGTCTTCTCTGTACCCAGGT 900
Qy 922 AGCAGACCTCGAAACTTGGATTTCAAGTCTGATTTTGGATTTCACTGGCTCTACTTTA 981
Db 901 AGCAGACCTCGAAACTTGGATTTCAAGTCTGATTTTGGATTTCACTGGCTCTACTTTA 960
Qy 982 TTTGTCTAATGCAAGACGATTAACATCTACATCTGTTTAAATATGACCTGGGTGAAGACTTCT 1041
Db 961 TTTGTCTAATGCAAGACGATTAACATCTACATCTGTTTAAATATGACCTGGGTGAAGACTTCT 1020
Qy 1042 CCAGTGGCTATTTCAATGGACACCAAGACTCTACCTTTTATGTAATCCAGCCCTTAGT 1101
Db 1021 CCAGTGGCTATTTCAATGGACACCAAGACTCTACCTTTTATGTAATCCAGCCCTTAGT 1080
Qy 1102 CCAGATGACAGCTTTTAGTCACTGGCTCAAGTGATGAAGCTGCCTACATATGGAGGTC 1161
Db 1081 CCAGATGACAGCTTTTAGTCACTGGCTCAAGTGATGAAGCTGCCTACATATGGAGGTC 1140
Qy 1162 TCCACACCTTGGCAACCTCTACTGTGCTCTGGGTCACTTCAAGAGGTCACGTCTGTG 1221
Db 1141 TCCACACCTTGGCAACCTCTACTGTGCTCTGGGTCACTTCAAGAGGTCACGTCTGTG 1200
Qy 1222 TGGTGTGCTCACTGACTTCAAAAGATTTGCTACCTTTTATGTAATCCAGCCCTTAGT 1281
Db 1201 TGGTGTGCTCACTGACTTCAAAAGATTTGCTACCTTTTATGTAATCCAGCCCTTAGT 1260
Qy 1282 ATCTGGCGCTTGAATAGAGCTTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCAGGTC 1341
Db 1261 ATCTGGCGCTTGAATAGAGCTTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCAGGTC 1320
Qy 1342 GGTGGGCTCTCAGAGAGAAAGAGTCAAGACCTTGGCTAGTACAGATTAACAGATGAC 1401
Db 1321 GGTGGGCTCTCAGAGAGAAAGAGTCAAGACCTTGGCTAGTACAGATTAACAGATGAC 1380
Qy 1402 CAGAGTACTTCTGCCAAGCCCGGCTAAGTGCAATCCATCCATTTCTTCCCGTCA 1461
Db 1381 CAGAGTACTTCTGCCAAGCCCGGCTAAGTGCAATCCATCCATTTCTTCCCGTCA 1440
Qy 1462 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACTCCCTCTCTCTTCCAAATACTCTACG 1521
Db 1441 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACTCCCTCTCTCTTCCAAATACTCTACG 1500
Qy 1522 TTTCTCTATTAAACCTCTCTGCAAGGCCCGGTCTCCCATCAAGAGAGGCTCTGTC 1581
Db 1501 TTTCTCTATTAAACCTCTCTGCAAGGCCCGGTCTCCCATCAAGAGAGGCTCTGTC 1560
Qy 1582 TCTCCGCTCTCTCCAGGCCACTTCTCATCTTTCAAGATGCGATTAGAACTGGGTGACC 1641
Db 1561 TCTCCGCTCTCTCCAGGCCACTTCTCATCTTTCAAGATGCGATTAGAACTGGGTGACC 1620
Qy 1642 CGAACACCTTCTCATCACCCACTCTCCACCTGCTTCCGAGACCAAGATCATGCT 1701
Db 1621 CGAACACCTTCTCATCACCCACTCTCCACCTGCTTCCGAGACCAAGATCATGCT 1680
Qy 1702 CCAGAGAAAGCCCTTATCTCTGTGAGCCAGAGTCTATCCCAAGCAGAGCTTGTCTCTGAG 1761
Db 1681 CCAGAGAAAGCCCTTATCTCTGTGAGCCAGAGTCTATCCCAAGCAGAGCTTGTCTCTGAG 1740
Qy 1762 TCTAGAAATAGATAAGAGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1821
Db 1741 TCTAGAAATAGATAAGAGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1800
Qy 1822 TGTGTGAAGATTGTAACCTGTGTGACTGAGCTTGTATGGCCCAAGTTCGAAATCTTCAATTG 1881
Db 1801 TGTGTGAAGATTGTAACCTGTGTGACTGAGCTTGTATGGCCCAAGTTCGAAATCTTCAATTG 1860

Qy 1882 GATCTGTGCTGCCCTTGTGTTAAACAGGAGACCTTAGTAGGACTCTCTAGTCTCTTACC 1941
Db 1861 GATCTGTGCTGCCCTTGTGTTAAACAGGAGACCTTAGTAGGACTCTCTAGTCTCTTACC 1920
Qy 1942 AAATCAAGCAAAATTTGAAGGAGCTGTGTACCAATATCTCAGAGCCTCGCTCTCTATCAGT 2001
Db 1921 AAATCAAGCAAAATTTGAAGGAGCTGTGTACCAATATCTCAGAGCCTCGCTCTCTATCAGT 1980
Qy 2002 CCGTATGCTTCCAGAAAGCTGTGGAAACGCTACCTCTTCTTTTGAACCTTGTGTGGAGAGGG 2061
Db 1981 CCGTATGCTTCCAGAAAGCTGTGGAAACGCTACCTCTTCTTTTGAACCTTGTGTGGAGAGGG 2040
Qy 2062 TCTGAATGCTAGGCAAGAGATAGTTCGCCAGAGATAAACTGGTCTTGGCCATG 2121
Db 2041 TCTGAATGCTAGGCAAGAGATAGTTCGCCAGAGATAAACTGGTCTTGGCCATG 2100
Qy 2122 GCAGCCAAACGGAAGGCTGAGAATCCATCTCCAGAACTCGCTCATCCAGACACCCCAAT 2181
Db 2101 GCAGCCAAACGGAAGGCTGAGAATCCATCTCCAGAACTCGCTCATCCAGACACCCCAAT 2160
Qy 2182 TCCAGGAGACAGCGGAAAGACATTTGCCAAAGCCGGTCAACATCAACGCCAGTCTCCATG 2241
Db 2161 TCCAGGAGACAGCGGAAAGACATTTGCCAAAGCCGGTCAACATCAACGCCAGTCTCCATG 2220
Qy 2242 AGGAAATCTGCACATACTTCCATAGAAAGTCCAGGAGGACTTCTGTGCTCTGAAAC 2301
Db 2221 AGGAAATCTGCACATACTTCCATAGAAAGTCCAGGAGGACTTCTGTGCTCTGAAAC 2280
Qy 2302 TCACAGAAATTTATAGATTCTTAATCTGAGTGAGTTTACTGAGCTTTGGTCCACTAAAAACAG 2361
Db 2281 TCACAGAAATTTATAGATTCTTAATCTGAGTGAGTTTACTGAGCTTTGGTCCACTAAAAACAG 2340
Qy 2362 CTGAGCTTTGGTCCACTAAAAACAGATGAAATAAATAAAGAGTGAATCTATTAACCTCTGTC 2421
Db 2341 CTGAGCTTTGGTCCACTAAAAACAGATGAAATAAATAAAGAGTGAATCTATTAACCTCTGTC 2400
Qy 2422 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2481
Db 2401 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2460
Qy 2482 TCTGTGCTCTACTACCAATATGATATGAGCTTCCAGAGGATGAATGCTGTGTTTAAAT 2541
Db 2461 TCTGTGCTCTACTACCAATATGATATGAGCTTCCAGAGGATGAATGCTGTGTTTAAAT 2520
Qy 2542 TCATAAGTAAATTTGTCTCTAGCATTTTGAATGAATGATCTTTCACTTTTAAATTAAT 2601
Db 2521 TCATAAGTAAATTTGTCTCTAGCATTTTGAATGAATGATCTTTCACTTTTAAATTAAT 2580
Qy 2602 TCATCTTCTCTATAATATGATATGATATGAGCTTCCAGAGGCAAAACAAAGTCTCTGTTA 2661
Db 2581 TCATCTTCTCTATAATATGATATGAGCTTCCAGAGGCAAAACAAAGTCTCTGTTA 2640
Qy 2662 TCCTGAAACTTTCTATGCTCAGTGAAAGATATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
Db 2641 TCCTGAAACTTTCTATGCTCAGTGAAAGATATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
Qy 2722 GGCTGACTGAGAAATCTCTCTGCTGAAGACCCCTGGTCTGTGCTGCTCCAAACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCTCTCTGCTGAAGACCCCTGGTCTGTGCTGCTCCAAACATGTATA 2760
Qy 2782 ATTTTATTTGAATATACATTAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAATATACATTAATCTTTTCACTATG 2793

RESULT 7

US-10-393-590-85
; Sequence 85, Application US/10393590
; Publication No. US20030190656A1
; GENERAL INFORMATION:
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: BREAST CANCER PROGNASTIC PORTFOLIO

FILE REFERENCE: CDS 268 US NP
CURRENT APPLICATION NUMBER: US/10/393,590
CURRENT FILING DATE: 2003-03-21
PRIOR APPLICATION NUMBER: 60/368,789
PRIOR FILING DATE: 2002-03-29
NUMBER OF SEQ ID NOS: 100
SOFTWARE: PatentIn version 3.1
SEQ ID NO 85
LENGTH: 4221
TYPE: DNA
ORGANISM: human
us-10-393-590-85

Query Match 98.5%; Score 2789,8; DB 7; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	22	CGATAACGATTTGGTTGTGAGAGGCGCAACGTCGCAATTTCTGCTGAACTTGGAGCAAT	81
DB	1	CGATAACGATTTGGTTGTGAGAGGCGCAACGTCGCAATTTCTGCTGAACTTGGAGCAAT	60
QY	82	TCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTCAATTCGGTG	141
DB	61	TCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTCAATTCGGTG	120
QY	142	CTCGCGCAGCCCAAGCTTGGCGCTCTGAGAAATGGATGGTCTTCAACAATACCCCTTTCAA	201
DB	121	CTCGCGCAGCCCAAGCTTGGCGCTCTGAGAAATGGATGGTCTTCAACAATACCCCTTTCAA	180
QY	202	TCCCTTTCTGACTGTTATCAGTCAGTCAGTGGTAATGATGAACACACTTTCTTATGGAGAACA	261
DB	181	TCCCTTTCTGACTGTTATCAGTCAGTCAGTGGTAATGATGAACACACTTTCTTATGGAGAACA	240
QY	262	GGAGTCCAGTTCCTCTTTTGATGATACCTTCTCTCTGCTCCCAATATGGAACATGTA	321
DB	241	GGAGTCCAGTTCCTCTTTTGATGATACCTTCTCTCTGCTCCCAATATGGAACATGTA	300
QY	322	CTAGCAGTTTCCCAATGAAGAAGCTTTGTTGCAATTTGATATAACACAGATCACAAGATTC	381
DB	301	CTAGCAGTTTCCCAATGAAGAAGCTTTGTTGCAATTTGATATAACACAGATCACAAGATTC	360
QY	382	AGAAAGAGTTCCTTCAAGAAATGGATGGCTCACTGGAATGCCGTCTTTTGACCTGGCTGG	441
DB	361	AGAAAGAGTTCCTTCAAGAAATGGATGGCTCACTGGAATGCCGTCTTTTGACCTGGCTGG	420
QY	442	GTTCTGGTGAATTAACCTTTGTAACGACGAGGTGATCAACACGCCAAATTTTGGGAC	501
DB	421	GTTCTGGTGAATTAACCTTTGTAACGACGAGGTGATCAACACGCCAAATTTTGGGAC	480
QY	502	GTAAAGCTGGTGAAGCTGATTTGTAACGACGAGGTGATCAACGACGCTCAAGTCAGTT	561
DB	481	GTAAAGCTGGTGAAGCTGATTTGTAACGACGAGGTGATCAACGACGCTCAAGTCAGTT	540
QY	562	GCCTTTTCTAAGTTTGAAGAAGCTGATTTCTGTACGGGTGGAAGAGATGGCAACATTATG	621
DB	541	GCCTTTTCTAAGTTTGAAGAAGCTGATTTCTGTACGGGTGGAAGAGATGGCAACATTATG	600
QY	622	GTCTGGGATACGAGGTGCAACAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAGT	681
DB	601	GTCTGGGATACGAGGTGCAACAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAGT	660
QY	682	GGAGCTCACAAATACCTCAGACAGCAACCCCTTCAAAACCCCAAGAGAAACAGAAATCA	741
DB	661	GGAGCTCACAAATACCTCAGACAGCAACCCCTTCAAAACCCCAAGAGAAACAGAAATCA	720
QY	742	AAAGGACTTGTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTTCAAGAC	801
DB	721	AAAGGACTTGTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTTCAAGAC	780
QY	802	GAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAGTATGGGATTTACGT	861
DB	781	GAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAGTATGGGATTTACGT	840

QY	862	AAGAATTATATCTGCTTATCGACAAGAACCCATAGCATCCAAGTCTTCTCTGTACCCAGGT	921
DB	841	AAGAATTATCTGCTTATCGACAAGAACCCATAGCATCCAAGTCTTCTCTGTACCCAGGT	900
QY	922	AGCAGCACTCGAAAACCTTGGATATTCAAGTCTGATTTTGGATTCCACTGGCTCTACTTTA	981
DB	901	AGCAGCACTCGAAAACCTTGGATATTCAAGTCTGATTTTGGATTCCACTGGCTCTACTTTA	960
QY	982	TTTGTCTAATTGCACAGACGATACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT	1041
DB	961	TTTGTCTAATTGCACAGACGATACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT	1020
QY	1042	CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAAAATCCAGCCTTAGT	1101
DB	1021	CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAAAATCCAGCCTTAGT	1080
QY	1102	CCAGATGACCAAGTTTGTAGTGGCTCAAGTGAATGAAGCTGCTTACATATGGAAGGTC	1161
DB	1081	CCAGATGACCAAGTTTGTAGTGGCTCAAGTGAATGAAGCTGCTTACATATGGAAGGTC	1140
QY	1162	TCCACACCTGCGCAACCTCTACTGTGCTCTGGGTCACTTCTCAAGAGGTCAAGTCTGTG	1221
DB	1141	TCCACACCTGCGCAACCTCTACTGTGCTCTGGGTCACTTCTCAAGAGGTCAAGTCTGTG	1200
QY	1222	TGCTGTGTCTCATCTGACTTCAAAAGATTGCTACCTGTTCTGATGACAATAACATAAAA	1281
DB	1201	TGCTGTGTCTCATCTGACTTCAAAAGATTGCTACCTGTTCTGATGACAATAACATAAAA	1260
QY	1282	ATCTGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGAGGTGATAAACTTTCCACGGTG	1341
DB	1261	ATCTGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGAGGTGATAAACTTTCCACGGTG	1320
QY	1342	GGTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGGCTAGTAAACAGTAAACAGTAGC	1401
DB	1321	GGTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGGCTAGTAAACAGTAAACAGTAGC	1380
QY	1402	CAGAGTACTCTCTGCCAAGCCCCCAGGGTAAAGTGAATCCATCCAAATTTTCCCCGTCA	1461
DB	1381	CAGAGTACTCTCTGCCAAGCCCCCAGGGTAAAGTGAATCCATCCAAATTTTCCCCGTCA	1440
QY	1462	TCCGAGCTTGTGCCCCCAAGCTGTGTGGAGACTCCCTCTTCTTCCAAATCTCTCTACG	1521
DB	1441	TCCGAGCTTGTGCCCCCAAGCTGTGTGGAGACTCCCTCTTCTTCCAAATCTCTCTACG	1500
QY	1522	TTCTCTATTAAACCTCTCTGCCAGGCCCGCTCTCCCATCAACAGAGAGGCTCTGTTC	1581
DB	1501	TTCTCTATTAAACCTCTCTGCCAGGCCCGCTCTCCCATCAACAGAGAGGCTCTGTTC	1560
QY	1582	TCTCCGTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTCGATTAGAAATCTGGGTGACC	1641
DB	1561	TCTCCGTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTCGATTAGAAATCTGGGTGACC	1620
QY	1642	CGAAACACCTTCTCATCACACCCATCTCACTCCACTGCTTCCGAGACCAAGATCATGTCT	1701
DB	1621	CGAAACACCTTCTCATCACACCCATCTCACTCCACTGCTTCCGAGACCAAGATCATGTCT	1680
QY	1702	CCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGCTTGTCTCTAG	1761
DB	1681	CCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGCTTGTCTCTAG	1740
QY	1762	TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAAG	1821
DB	1741	TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAAG	1800
QY	1822	TGTGTGAAGAGTTGTAATCTGTGTGAGTGTGAGTGGCCAAAGTTGAAAATCTTCAATTTG	1881
DB	1801	TGTGTGAAGAGTTGTAATCTGTGTGAGTGTGAGTGGCCAAAGTTGAAAATCTTCAATTTG	1860
QY	1882	GATCTGTGCTGCTTGTCTGTTAACCGAGAGACCTTAGTAGGACTCTCTAGTGTCTTACC	1941
DB	1861	GATCTGTGCTGCTTGTCTGTTAACCGAGAGACCTTAGTAGGACTCTCTAGTGTCTTACC	1920
QY	1942	AAATCAAGCAAAATTTGAAGAGGCTGGTACCAGTATCTCAGAGGCTCCGCTCTCTATCAGT	2001


```
|||||
1921 AAATCAAGCAAAATTGAGAGAGCTGTACAGATATCTCAGAGCTCCGCTCTCTATCAGT 1980
QY CCGTATGCTTCAGAAAGCTGTGGACGCTACCTCTTCCTTTGAGAGCTTTGTGGAGAGGG 2061
Db CCGTATGCTTCAGAAAGCTGTGGACGCTACCTCTTCCTTTGAGAGCTTTGTGGAGAGGG 2040
QY TCTGAAATGTTAGGCAAGAGAAATAGTTCCCGAGAGAAATAAAACHTGGTTGTGGCCATG 2121
Db TCTGAAATGTTAGGCAAGAGAAATAGTTCCCGAGAGAAATAAAACHTGGTTGTGGCCATG 2100
QY GCAGCCAAACGGAAGGCTGAGAAATCACTCTCCACGAAGTCCGTCCTCCAGACACCCCAAT 2181
Db GCAGCCAAACGGAAGGCTGAGAAATCACTCTCCACGAAGTCCGTCCTCCAGACACCCCAAT 2160
QY TCCAGGAGACAGACGCGGAAAGACAATGGCCAAAGCCCGGTCAACATCAGCCCGAGCTCCATG 2241
Db TCCAGGAGACAGACGCGGAAAGACAATGGCCAAAGCCCGGTCAACATCAGCCCGAGCTCCATG 2220
QY AGGAAATCTGCACATACCTTCCATAGAAAGTCCAGAGAGACTTCTGTGGTCTGAAACAC 2301
Db AGGAAATCTGCACATACCTTCCATAGAAAGTCCAGAGAGACTTCTGTGGTCTGAAACAC 2280
QY TCAACAGAAATTATAGATTCTTAATCTCAGTGAGTTACTGAGCTTTGGTCCACTAAACACAG 2361
Db TCAACAGAAATTATAGATTCTTAATCTCAGTGAGTTACTGAGCTTTGGTCCACTAAACACAG 2340
QY CTGAGCTTTGGTCCACTAAACCAAGATGAAATAACAAGAGTGACTCTATAACTCTGGTC 2421
Db CTGAGCTTTGGTCCACTAAACCAAGATGAAATAACAAGAGTGACTCTATAACTCTGGTC 2400
QY TTTAAGAAAGCTGCCCTTTTCATTTTATAGACAAATCTTTTCAACGCTGAAATGTACTTAA 2481
Db TTTAAGAAAGCTGCCCTTTTCATTTTATAGACAAATCTTTTCAACGCTGAAATGTACTTAA 2460
QY TCTGGTTCTACTACCAATAATGTATATGACAGCTTCCGAGGAGTAAATGCTGTGTTAAAT 2541
Db TCTGGTTCTACTACCAATAATGTATATGACAGCTTCCGAGGAGTAAATGCTGTGTTAAAT 2520
QY TCATAAAGTAAATTTGTCACTCTAGCAATTTTGAATAGTCTTCACTTTTAAATATAT 2601
Db TCATAAAGTAAATTTGTCACTCTAGCAATTTTGAATAGTCTTCACTTTTAAATATAT 2580
QY TCATCTTCTATAAATAAGACATCCAGTTTCATGGAGGCAAAAAACAAGTTTCTTGTTA 2661
Db TCATCTTCTATAAATAAGACATCCAGTTTCATGGAGGCAAAAAACAAGTTTCTTGTTA 2640
QY TCCGGAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCAGCATGAGGCTGTGAA 2721
Db TCCGGAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCAGCATGAGGCTGTGAA 2700
QY GGCTGACTGAGAAATCTCTGCTGGAAGACCCCTGGTTCTGTTCTGCTCCAAATGTATA 2781
Db GGCTGACTGAGAAATCTCTGCTGGAAGACCCCTGGTTCTGTTCTGCTCCAAATGTATA 2760
QY ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db ATTTTATTTGAAATACATAATCTTTTCACTATG 2793
```

RESULT 8

```
US-10-393-567-85
; Sequence 85, Application US/10393567
; Publication No. US20030194733A1
; GENERAL INFORMATION:
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: CANCER DIAGNOSTIC PANEL
; FILE REFERENCE: CDS 269 US NP
; CURRENT APPLICATION NUMBER: US/10/393,567
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: 60/368,667
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 100
```

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 85

; LENGTH: 4221

; TYPE: DNA

; ORGANISM: human

US-10-393-567-85

Query Match 98.5%; Score 2789.8; DB 7; Length 4221;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 CGATAACGATTGTCTTCTGAGAGCGCAACGTCGGATTTCTGCTGGAACCTTTGGAGGCAAT 81

Db 1 CGATHAACGATTGTCTTCTGAGAGCGCAACGTCGGATTTCTGCTGGAACCTTTGGAGGCAAT 60

QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 141

Db 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 120

QY 142 CTCGCCACGCCAGCTTTGGGCTCTGAGAAATGGATGGTCTTACAAATACCTCTTCAA 201

Db 121 CTCGCCACGCCAGCTTTGGGCTCTGAGAAATGGATGGTCTTACAAATACCTCTTCAA 180

QY 202 TCCCTTCTGACTGGTTATCAGTGCACTGATGTAATGTAACACACTTCTTATGGAGAAACA 261

Db 181 TCCCTTCTGACTGGTTATCAGTGCACTGATGTAATGTAACACACTTCTTATGGAGAAACA 240

QY 262 GGAGTCCCAGTTCTCTCTTTGGATGTACCTTCTTCTGCTCCCAATATGGAACATGTA 321

Db 241 GGAGTCCCAGTTCTCTCTTTGGATGTACCTTCTTCTGCTCCCAATATGGAACATGTA 300

QY 322 CTAGCAGTTGCCAATGAAAGGCTTTCTGATTTGTAACACAGAAATCAAAAGTTTC 381

Db 301 CTAGCAGTTGCCAATGAAAGGCTTTCTGATTTGTAACACAGAAATCAAAAGTTTC 360

QY 382 AGAAGAAAGTGTCTCAAGAAATGGAATGCTCACTGGAATCGCTTTTGACCTGCGCTGG 441

Db 361 AGAAGAAAGTGTCTCAAGAAATGGAATGGAATGCTCACTGGAATCGCTTTTGACCTGCGCTGG 420

QY 442 GTTCTCGTGAACTTAAACTTTGTTACAGCAGCAGGTGATCAAAACAGCCAAATTTGGGAC 501

Db 421 GTTCTCGTGAACTTAAACTTTGTTACAGCAGCAGGTGATCAAAACAGCCAAATTTGGGAC 480

QY 502 GTAAAAGCTGTGAGCTGATTTGGAAACATGCAAAAGTCAATGAGCCTCAAGTCAGTT 561

Db 481 GTAAAAGCTGTGAGCTGATTTGGAAACATGCAAAAGTCAATGAGCCTCAAGTCAGTT 540

QY 562 GCCTTTTCTAAGTTTGAGAAAGTCTGATTTCTGACGGTGGAAAGATGGCAACATTATG 621

Db 541 GCCTTTTCTAAGTTTGAGAAAGTCTGATTTCTGACGGTGGAAAGATGGCAACATTATG 600

QY 622 GTCTGGGATACAGGTGCAACAAAAAGATGGTTTTTATAGGCAAGTGAATCAAAATCAGT 681

Db 601 GTCTGGGATACAGGTGCAACAAAAAGATGGTTTTTATAGGCAAGTGAATCAAAATCAGT 660

QY 682 GGAGCTCACAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGNAGNAGAGATTCA 741

Db 661 GGAGCTCACAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGNAGNAGAGATTCA 720

QY 742 AAAGGACTTTGCTCTTCTGTGGATTTCCAGCAAAAGTGTACTGTGCTCTCTTTCAAGAC 801

Db 721 AAAGGACTTTGCTCTTCTGTGGATTTCCAGCAAAAGTGTACTGTGCTCTCTTTCAAGAC 780

QY 802 GAGAATACTTACTGCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 861

Db 781 GAGAATACTTACTGCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 840

QY 862 AAGAATTATCTGCTTATCGACAAGAACCCATAGCATCCAGTCCAGTCTTCTGCTACCCAGT 921

Db 841 AAGAATTATCTGCTTATCGACAAGAACCCATAGCATCCAGTCTTCTGCTACCCAGT 900

QY 922 AGCAGCACTCGAAAAAATTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981

Query Match		98.5%	Score 2789.8;	DB 7;	Length 4221;	
Best Local Similarity		99.9%;	Pred. No. 0;			
Matches 2791;		Conservative	0;	Mismatches	2;	Indels 0; Gaps 0;
Q	22	CGATAACGATTGTGTGTGAGAGCGCAACGTCGGATTTCTGCTGAACCTTGAGGCAATT	81			
Q	1	CGATAACGATTGTGTGTGAGAGCGCAACGTCGGATTTCTGCTGAACCTTGAGGCAATT	60			
Q	82	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGAGCCCTGATGCTCTTCAATTCGGTG	141			
Q	61	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGAGCCCTGATGCTCTTCAATTCGGTG	120			
Q	142	CTCCGCCAGCCCGCTTGGCGTCTCGAAGAAATGATGGTCTTCCACAATACCCCTCTTCAA	201			
Q	121	CTCCGCCAGCCCGCTTGGCGTCTCGAAGAAATGATGGTCTTCCACAATACCCCTCTTCAA	180			
Q	202	TCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA	261			
Q	181	TCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA	240			
Q	262	GGAGTCCCAGTTCCTCTTTTGGATGTACCTTCTCTTCTGCTCCCAATATGGAACATGTA	321			
Q	241	GGAGTCCCAGTTCCTCTCTTTTGGATGTACCTTCTCTTCTGCTCCCAATATGGAACATGTA	300			
Q	322	CTAGCAGTTGCCAATGAAGAAGGCTTTGTTTGCATTTGTATACACAGAAATCACAAAGTTTC	381			
Q	301	CTAGCAGTTGCCAATGAAGAAGGCTTTGTTTGCATTTGTATACACAGAAATCACAAAGTTTC	360			
Q	382	AGAAAGAGTGTCTCAAGAATGGATGGCTCACTGGAATGCCGTCTTTGACCTGGCCTGG	441			
Q	361	AGAAAGAGTGTCTCAAGAATGGATGGCTCACTGGAATGCCGTCTTTGACCTGGCCTGG	420			
Q	442	GTTCTCTGTGAACTTAACTTGTATACAGCAGAGTGATCAACAGCCAAATTTTGGGAC	501			
Q	421	GTTCTCTGTGAACTTAACTTGTATACAGCAGAGTGATCAACAGCCAAATTTTGGGAC	480			
Q	502	GTAAAAGCTGTGTGAGTGTGAACATGCAAAAGTCAATCAATGCAGCCTCAAGTCAGTT	561			
Q	481	GTAAAAGCTGTGTGAGTGTGAACATGCAAAAGTCAATCAATGCAGCCTCAAGTCAGTT	540			
Q	562	GCCTTTCTTAAGTTTGAGAAAGCTGTATCTGTACGGGTGGAAGAGATGGCAACATTATG	621			
Q	541	GCCTTTCTTAAGTTTGAGAAAGCTGTATCTGTACGGGTGGAAGAGATGGCAACATTATG	600			
Q	622	GTCTGGGATACAGGTGCAACAAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT	681			
Q	601	GTCTGGGATACAGGTGCAACAAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT	660			
Q	682	GGAGTCTCAATATCCTCAGACAGCAAAACCCCTTCAAAACCCAGAGAAACAGAAATTC	741			
Q	661	GGAGTCTCAATATCCTCAGACAGCAAAACCCCTTCAAAACCCAGAGAAACAGAAATTC	720			
Q	742	AAAGGACTTGCTCTCTGTGATTTCCAGCAAGTGTACTGTGCTCTTCTTCAAGAC	801			
Q	721	AAAGGACTTGCTCTCTGTGATTTCCAGCAAGTGTACTGTGCTCTTCTTCAAGAC	780			
Q	802	GAGAAATACCTTGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT	861			
Q	781	GAGAAATACCTTGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT	840			
Q	862	AAGAATATATGCTTTATCGAACAAGACCCATAGCATCAAGTCTTTCTGTATCCCAAGGT	921			
Q	841	AAGAATATATGCTTTATCGAACAAGACCCATAGCATCAAGTCTTTCTGTATCCCAAGGT	900			
Q	922	AGCAGCACTCGAAAACTTTGGATATTCAAGTCTGATTTTGGGATTTCCACTGGCTCTACTTTA	981			
Q	901	AGCAGCACTCGAAAACTTTGGATATTCAAGTCTGATTTTGGGATTTCCACTGGCTCTACTTTA	960			
Q	982	TTTGTCTAAATTCAGACAGATACATCTACATGTTTAAATATGACTGGGTGGAAGACTTCT	1041			
Q	961	TTTGTCTAAATTCAGACAGATACATCTACATGTTTAAATATGACTGGGTGGAAGACTTCT	1020			

Qy	1042	CCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT	1101
Db	1021	CCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT	1080
Qy	1102	CCAGATGACCAAGTTTATGTAGTCAGTGGCTCAAGTGATGAAGCTGCCTACATATGGAAGTTC	1161
Db	1081	CCAGATGACCAAGTTTATGTAGTCAGTGGCTCAAGTGATGAAGCTGCCTACATATGGAAGTTC	1140
Qy	1162	TCCACACCTGGGCAACCTCTTACTGTGCTCTGGGTGATTTCTCAAGAGGTCAAGTCTGTG	1221
Db	1141	TCCACACCTGGGCAACCTCTTACTGTGCTCTGGGTGATTTCTCAAGAGGTCAAGTCTGTG	1200
Qy	1222	TGCTGGTGTCCATCTGACTTCCAAAGATTCCTACTCTGTTCTGATGACAAATACATAAAA	1281
Db	1201	TGCTGGTGTCCATCTGACTTCCAAAGATTCCTACTCTGTTCTGATGACAAATACATAAAA	1260
Qy	1282	ATCTGGGCTTTGAATAGAGGCTTAGAGAGAAACCCAGAGGTGATAAATTTTCCACGGTG	1341
Db	1261	ATCTGGGCTTTGAATAGAGGCTTAGAGAGAAACCCAGAGGTGATAAATTTTCCACGGTG	1320
Qy	1342	GGTTGGGCTCTCAGAGAAACCAAGAGTCAAGACCTGGCCCTAGTAAACAGTAAACGATGAC	1401
Db	1321	GGTTGGGCTCTCAGAGAAACCAAGAGTCAAGACCTGGCCCTAGTAAACAGTAAACGATGAC	1380
Qy	1402	CAGAGTACTCTGCAAGACCCCGAGGTAAAGTGCAATCCATCCAATTTTCCCGTCA	1461
Db	1381	CAGAGTACTCTGCAAGACCCCGAGGTAAAGTGCAATCCATCCAATTTTCCCGTCA	1440
Qy	1462	TCCGAGCTTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATCTCCTACG	1521
Db	1441	TCCGAGCTTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATCTCCTACG	1500
Qy	1522	TTCTCTATTAACACTCTCTCTGCCAAGGCCCGGTCTCCCATCAACAGAAAGGCTCTGTCT	1581
Db	1501	TTCTCTATTAACACTCTCTCTGCCAAGGCCCGGTCTCCCATCAACAGAAAGGCTCTGTCT	1560
Qy	1582	TCCTCCGCTCTCTCCCAAGCCACTTCACTTTTCAAGATGTCGATTAGAAATCTGGGTGACC	1641
Db	1561	TCCTCCGCTCTCTCCCAAGCCACTTCACTTTTCAAGATGTCGATTAGAAATCTGGGTGACC	1620
Qy	1642	CGAACACCTTCTCTCATCAACCACTCACTCCACCTGCTTCCGGAGACCAAGATCATGTCT	1701
Db	1621	CGAACACCTTCTCTCATCAACCACTCACTCCACCTGCTTCCGGAGACCAAGATCATGTCT	1680
Qy	1702	CGAGAAAGCCCTTATTCCTGTGAGCCAGAACTCATCCCAAGCAGAGGCTTGGCTCTGAG	1761
Db	1681	CGAGAAAGCCCTTATTCCTGTGAGCCAGAACTCATCCCAAGCAGAGGCTTGGCTCTGAG	1740
Qy	1762	TCTAGAAATAGATTAAGAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG	1821
Db	1741	TCTAGAAATAGATTAAGAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG	1800
Qy	1822	TGTGTGAAGAGTGTGTAAGTGTGACTGTGAGCTTGAAGCCCAAGTTGAAATCTTCAATTTG	1881
Db	1801	TGTGTGAAGAGTGTGTAAGTGTGACTGTGAGCTTGAAGCCCAAGTTGAAATCTTCAATTTG	1860
Qy	1882	GATCTGTGCTGCTGTCTGTGTAACCAAGAGACCTTAGTAGAGACTCTCTAGGTCTTACC	1941
Db	1861	GATCTGTGCTGCTGTCTGTGTAACCAAGAGACCTTAGTAGAGACTCTCTAGGTCTTACC	1920
Qy	1942	AAATCAAGCAAAATTTGAAGGAGCTGTTACCAAGTATCTCAGAGCCCTCCCTCTATCAGT	2001
Db	1921	AAATCAAGCAAAATTTGAAGGAGCTGTTACCAAGTATCTCAGAGCCCTCCCTCTATCAGT	1980
Qy	2002	CCGTATGCTTCCAGAAAGCTGTGGAAACGCTACCTCTCTTCTTGGAGACCTTGTGGAGAGGG	2061
Db	1981	CCGTATGCTTCCAGAAAGCTGTGGAAACGCTACCTCTCTTCTTGGAGACCTTGTGGAGAGGG	2040
Qy	2062	TCTGAAATGTTAGGCAAGAGATAGTTCCCGAGAGATATAAACTGGTGTGGTGGCCATG	2121
Db	2041	TCTGAAATGTTAGGCAAGAGATAGTTCCCGAGAGATATAAACTGGTGTGGTGGCCATG	2100
Qy	2122	GCAGCCAAACCGAAGGCTGAGAAATCCATCTCCACGAAGTCCGTATCCAGACACACCAAT	2181

Db 2101 GCAGCCAAACGGAAGCTGAGATCCATCTCCAGGAAGTCGGTCATCCAGACACCCCAAT 2160
Qy 2182 TCAGAGAGACAGAGCGGAAAGACATTCGCCAGCCCGGTCCATCAGCCCGAGCTCCATG 2241
Db 2161 TCAGAGAGACAGAGCGGAAAGACATTCGCCAGCCCGGTCCATCAGCCCGAGCTCCATG 2220
Qy 2242 AGGAAATCTGCACATACCTCCATAGAAAGTCCAGAGGACTTCCTGTGCTCTGAACAC 2301
Db 2221 AGGAAATCTGCACATACCTCCATAGAAAGTCCAGAGGACTTCCTGTGCTCTGAACAC 2280
Qy 2302 TCAACAGAAATATPAGATTCCTAATCTCTGAGTGAGTACTTGAGCTTTGGTCCACTAAACAAAG 2361
Db 2281 TCAACAGAAATATPAGATTCCTAATCTCTGAGTGAGTACTTGAGCTTTGGTCCACTAAACAAAG 2340
Qy 2362 CTGAGCTTTGGTCCACTAAACAAAGATGAAATATACAGAGTGACTCTATTAATCTGGTC 2421
Db 2341 CTGAGCTTTGGTCCACTAAACAAAGATGAAATATACAGAGTGACTCTATTAATCTGGTC 2400
Qy 2422 TTTAAGAAAGCTGCCCTTTTCAATTTTAGACAAATCTTTTCAACGCTGAAATGTACTTAA 2481
Db 2401 TTTAAGAAAGCTGCCCTTTTCAATTTTAGACAAATCTTTTCAACGCTGAAATGTACTTAA 2460
Qy 2482 TCTGGTTCTACTACCAATATGTATATGCAGCTTCCGAGGATGAATGCTGTGTTTAAAT 2541
Db 2461 TCTGGTTCTACTACCAATATGTATATGCAGCTTCCGAGGATGAATGCTGTGTTTAAAT 2520
Qy 2542 TCATAAGTAAATTTGTGACTCTAGCAATTTGAATGAATAGTCTTCACTTTTAAATAT 2601
Db 2521 TCATAAGTAAATTTGTGACTCTAGCAATTTGAATGAATAGTCTTCACTTTTAAATAT 2580
Qy 2602 TCATCTTCTCTATAATAATGACATCCAGTTCATGAGGCAAAACAAAGTTCTTGTTA 2661
Db 2581 TCATCTTCTCTATAATAATGACATCCAGTTCATGAGGCAAAACAAAGTTCTTGTTA 2640
Qy 2662 TCCGAAACTTTCTATGCTCAGTGGAAAGTATCTGCAGCCACAGCATGAGGCTGTGAA 2721
Db 2641 TCCGAAACTTTCTATGCTCAGTGGAAAGTATCTGCAGCCACAGCATGAGGCTGTGAA 2700
Qy 2722 GGTGACTGAGAAATCCTCTGCTGAAAGACCCCTGGTCTGTTCTGCTCCAAATGTATA 2781
Db 2701 GGTGACTGAGAAATCCTCTGCTGAAAGACCCCTGGTCTGTTCTGCTCCAAATGTATA 2760
Qy 2782 ATTTATTTGAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTATTTGAATACATAATCTTTTCACTATG 2793

RESULT 10
US-10-172-118-1645
; Sequence 1645, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linesley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1645
; LENGTH: 4221
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: NM_016448
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1645
Query Match 98.5%; Score 2789.8; DB 7; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 22 CGATAACGATTTGCTCTGAGAGCGCAACGTCGATTTCTGCTGAACTTCGAGGCAT 81
Db 1 CGATAACGATTTGCTCTGAGAGCGCAACGTCGATTTCTGCTGAACTTCGAGGCAT 60
Qy 82 TCTAGACATTTCTCTCAGCTGAGGCTTTTCTCCGACCTCATGCTCTTCAATTCGGTG 141
Db 61 TCTAGACATTTCTCTCAGCTGAGGCTTTTCTCCGACCTCATGCTCTTCAATTCGGTG 120
Qy 142 CTCGCCACGCCACGCTTGGGCTCTGAGAAATGATGGTCTTCAATACCTCTTCAA 201
Db 121 CTCGCCACGCCACGCTTGGGCTCTGAGAAATGATGGTCTTCAATACCTCTTCAA 180
Qy 202 TCCCTTCTGACTGGTTATCAGTGCAGTGGTATGATGAACACACTTCTTATGGAAACA 261
Db 181 TCCCTTCTGACTGGTTATCAGTGCAGTGGTATGATGAACACACTTCTTATGGAAACA 240
Qy 262 GGAGTCCCAGTTCCTCTTTTGGATGTACCTTCTCTCTGCTCCCAATATGAAACATGTA 321
Db 241 GGAGTCCCAGTTCCTCTTTTGGATGTACCTTCTCTCTGCTCCCAATATGAAACATGTA 300
Qy 322 CTAGCAGTTCGCAATGAAAGGCTTTGTCGATTTGATATACACAGAAATCAAAAGTTTC 381
Db 301 CTAGCAGTTCGCAATGAAAGGCTTTGTCGATTTGATATACACAGAAATCAAAAGTTTC 360
Qy 382 AGAAAGATGCTTCAAAGATGGATGGCTCACTGGAATGCCGCTTTTGACCTGGCTGG 441
Db 361 AGAAAGATGCTTCAAAGATGGATGGCTCACTGGAATGCCGCTTTTGACCTGGCTGG 420
Qy 442 GTTCTGCTGAACTTTAAACTTGTACAGCAGCAGGTGATCAAAACAGCAGCAATTTTGGGAC 501
Db 421 GTTCTGCTGAACTTTAAACTTGTACAGCAGCAGGTGATCAAAACAGCAGCAATTTTGGGAC 480
Qy 502 GTAAAAGCTGGTGAAGTCAATGGAAATGCAAGGTGATCAATGAGCTCAAGTCAGTT 561
Db 481 GTAAAAGCTGGTGAAGTCAATGGAAATGCAAGGTGATCAATGAGCTCAAGTCAGTT 540
Qy 562 GCTTTTCTAAGTTTGAGAAAGCTGATTTCTGACGGGTGGAGAGATGGCAACATTATG 621
Db 541 GCTTTTCTAAGTTTGAGAAAGCTGATTTCTGACGGGTGGAGAGATGGCAACATTATG 600
Qy 622 GTCTGGGATACCAGGTGCAACAAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAGT 681
Db 601 GTCTGGGATACCAGGTGCAACAAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAGT 660
Qy 682 GGAGCTCAAAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATCA 741
Db 661 GGAGCTCAAAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATCA 720
Qy 742 AAAGGACTTGTCTCTGAGATTTCCAGCAAGTGTACTGTGCTCTCTTCAAGAC 801
Db 721 AAAGGACTTGTCTCTGAGATTTCCAGCAAGTGTACTGTGCTCTCTTCAAGAC 780
Qy 802 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAATCAAAATATGGGATTTACGT 861
Db 781 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAATCAAAATATGGGATTTACGT 840
Qy 862 AAGAAATTATCTGCTTATCGAACAAGACCCATAGCATCCAAGTCTTTCTGTACCAGGT 921
Db 841 AAGAAATTATCTGCTTATCGAACAAGACCCATAGCATCCAAGTCTTTCTGTACCAGGT 900
Qy 922 AGCAGACTCGAAACTTGGATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
Db 901 AGCAGACTCGAAACTTGGATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
Qy 982 TTTGCTAATTGCACAGACGATAACATCTACATGTTTTAATATGACTGGGTTGAAGACTTCT 1041

||||| 961 TTTGCTAATTGACAGAGATTAACATCTACATGTTTATATGACTGGGTGAGACTTCT 1020
||||| 1042 CCAGTGGCTATTTTCAATGAGACACAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1101
||||| 1021 CCAGTGGCTATTTTCAATGAGACACAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1080
||||| 1102 CCAGATGACAGTATTTTAGTCACTGAGTCAAGTGAATGAAGCTGCCTACATATGGAAGGTC 1161
||||| 1081 CCAGATGACAGTATTTTAGTCACTGAGTCAAGTGAATGAAGCTGCCTACATATGGAAGGTC 1140
||||| 1162 TCCACACCTGCGCAACCTCTACTGTGCTCCTGGGTCAATTTCTCAAGAGGTCACTGCTGTG 1221
||||| 1141 TCCACACCTGCGCAACCTCTACTGTGCTCCTGGGTCAATTTCTCAAGAGGTCACTGCTGTG 1200
||||| 1222 TGCTGGTGTCCATCTGACTTCTCAAAAGATTTGCTACCTGTTCTGTATGACAAATACACTAAAA 1281
||||| 1201 TGCTGGTGTCCATCTGACTTCTCAAAAGATTTGCTACCTGTTCTGTATGACAAATACACTAAAA 1260
||||| 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACAGAGAGGTGATAACTTTCCACGGTG 1341
||||| 1261 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACAGAGAGGTGATAACTTTCCACGGTG 1320
||||| 1342 GGTGTGGGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCTAGTAAACAGATACAGAGTAGC 1401
||||| 1321 GGTGTGGGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCTAGTAAACAGATACAGAGTAGC 1380
||||| 1402 CAGAGTACTCTCTGCCAAGCCCCCAGGGTAAAGTGCAATCCATCAAAATTTCTTCCCGGTCA 1461
||||| 1381 CAGAGTACTCTCTGCCAAGCCCCCAGGGTAAAGTGCAATCCATCAAAATTTCTTCCCGGTCA 1440
||||| 1462 TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCTAGG 1521
||||| 1441 TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCTAGG 1500
||||| 1522 TTCTCTATTAAACCTCTCTGCCAAGCCCCCAGGGTAAAGTGCAATCCATCAAAATTTCTTCCCGGTCA 1581
||||| 1501 TTCTCTATTAAACCTCTCTGCCAAGCCCCCAGGGTAAAGTGCAATCCATCAAAATTTCTTCCCGGTCA 1560
||||| 1582 TCCTCGCTCTCTCCCAAGCACCTTCTCATCTTTCAAGATGTGATAGAACTGGGTGACC 1641
||||| 1561 TCCTCGCTCTCTCCCAAGCACCTTCTCATCTTTCAAGATGTGATAGAACTGGGTGACC 1620
||||| 1642 CGAACACCTTCTCTATCAACCCATCACTCACTGCTTGGAGACCAAGATCATGTCT 1701
||||| 1621 CGAACACCTTCTCTATCAACCCATCACTCACTGCTTGGAGACCAAGATCATGTCT 1680
||||| 1702 CCGAGAAAGCCCTTATTCTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTCGAG 1761
||||| 1681 CCGAGAAAGCCCTTATTCTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTCGAG 1740
||||| 1762 TCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1821
||||| 1741 TCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1800
||||| 1822 TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTGATGGCCAAAGTTGAAATCTTCAATTG 1881
||||| 1801 TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTGATGGCCAAAGTTGAAATCTTCAATTG 1860
||||| 1882 GATCTGTGCTGCTTGTCTGTAACCAAGGAGACCTTTAGTAAGGACTCTCTAGGCTCTTACC 1941
||||| 1861 GATCTGTGCTGCTTGTCTGTAACCAAGGAGACCTTTAGTAAGGACTCTCTAGGCTCTTACC 1920
||||| 1942 AAATCAAGCAAAATTTGAAGAGCTGTGTACAGTATCTCAGAGCCTCCGCTCTCTATCAGT 2001
||||| 1921 AAATCAAGCAAAATTTGAAGAGCTGTGTACAGTATCTCAGAGCCTCCGCTCTCTATCAGT 1980
||||| 2002 CCGTATGCTTTCAGAAAGCTGTGGAACGCTACCTCTTCTTTCAGACCTTGTGGAGAGGG 2061
||||| 1981 CCGTATGCTTTCAGAAAGCTGTGGAACGCTACCTCTTCTTTCAGACCTTGTGGAGAGGG 2040
||||| 2062 TCTGAAATCGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACTGGTGTGTCCTATG 2121

Db 2041 TCTGAAATGTTAGCAAAAGAGAAATAGTTCCCGAGAGAAATAAAAACTGTTGTTGGCCATG 2100
Qy 2122 GCAGCCAAACCGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTATCCAGACACACCCCAAT 2181
Db 2101 GCAGCCAAACCGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTATCCAGACACACCCCAAT 2160
Qy 2182 TCCAGGAGACAGAGCGGGAAGACACATTCGCAAGCCCGGTCCACATCAGCCCGAGCTCCCATG 2241
Db 2161 TCCAGGAGACAGAGCGGGAAGACATTCGCAAGCCCGGTCCACATCAGCCCGAGCTCCCATG 2220
Qy 2242 AGGAAATCTCCACATACATCTCCATAGAAAGTCCCGAGGAGACTTCTGTGGTCTCTGAAACAC 2301
Db 2221 AGGAAATCTCCACATACATCTCCATAGAAAGTCCCGAGGAGACTTCTGTGGTCTCTGAAACAC 2280
Qy 2302 TCAACAGAAATATAGATTTCTAATCTGAGTGAATTTCTGAGCTTTGAGCTTCCATCAAAACAG 2361
Db 2281 TCAACAGAAATATAGATTTCTAATCTGAGTGAATTTCTGAGCTTTGGTCCATCAAAACAG 2340
Qy 2362 CTGAGCTTTGGTCCACTAAAAACAGATGAAATAACAAGAGTGAATCTATAACTCTGTC 2421
Db 2341 CTGAGCTTTGGTCCACTAAAAACAGATGAAATAACAAGAGTGAATCTATAACTCTGTC 2400
Qy 2422 TTTAAGAAAGCTGCTTTTCAATTTTAGACAAAATCTTTTCAACGCTGAAATGTACCTAA 2481
Db 2401 TTTAAGAAAGCTGCTTTTCAATTTTAGACAAAATCTTTTCAACGCTGAAATGTACCTAA 2460
Qy 2482 TCTGGTCTCTACATCAATATGATATGAGCTTCCGAGGATGAATGCTGTGTTAAATTT 2541
Db 2461 TCTGGTCTCTACATCAATATGATATGAGCTTCCGAGGATGAATGCTGTGTTAAATTT 2520
Qy 2542 TCATAAAGTAAATTTGTACACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2601
Db 2521 TCATAAAGTAAATTTGTACACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2580
Qy 2602 TCATCTTCTATATATATGACATCCAGTTCATGGAGGCAAAAACAAAGTTTCTTGTTA 2661
Db 2581 TCATCTTCTATATATATGACATCCAGTTCATGGAGGCAAAAACAAAGTTTCTTGTTA 2640
Qy 2662 TCCTGAAACTTTTCTATGCTAGTGGAAGATCTGCCAGGCCACAGCATGAGGCTGTGAA 2721
Db 2641 TCCTGAAACTTTTCTATGCTAGTGGAAGATCTGCCAGGCCACAGCATGAGGCTGTGAA 2700
Qy 2722 GGCTGACTGAGAAATCCCTCTGCTGAAGACCCCTGCTGCTTCTGCTCCCAACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCCCTCTGCTGAAGACCCCTGCTGCTTCTGCTCCCAACATGTATA 2760
Qy 2782 ATTTATTTGAAATACATAAATCTTTTCACTATG 2814
Db 2761 ATTTATTTGAAATACATAAATCTTTTCACTATG 2793

RESULT 11

US-10-342-887-1645
; Sequence 1645, Application US/10342387
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118

;; PRIOR FILING DATE: 2002-06-14
;; NUMBER OF SEQ ID NOS: 2699
;; SEQ ID NO 1645
;; LENGTH: 4221
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-342-887-1645

Query Match 98.5%; Score 2789.8; DB 8; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	22	CGATAACGATTTGTGTTGAGAGGCGCAACGTCGCAATTTCTGCTGAACCTTGGAGGCATT	81
Db	1	CGATAACGATTTGTGTTGAGAGGCGCAACGTCGCAATTTCTGCTGAACCTTGGAGGCATT	60
Qy	82	TCTACGACTTTCTCTCAGCTGAGGCTTTTCTCGACCCCTGATGCTCTCAATTTGGTG	141
Db	61	TCTACGACTTTCTCTCAGCTGAGGCTTTTCTCGACCCCTGATGCTCTCAATTTGGTG	120
Qy	142	CTCGGCGAGCCCGAGCTTGGCGTCTGAGAAATGGATGGTCTTTCACAAATACCCCTCTTCAA	201
Db	121	CTCGGCGAGCCCGAGCTTGGCGTCTGAGAAATGGATGGTCTTTCACAAATACCCCTCTTCAA	180
Qy	202	TCCTTTCTGACTGTTATCAGTGCAGTGGTAAATGATGAAACACTTCTTTATGAGAAACA	261
Db	181	TCCTTTCTGACTGTTATCAGTGCAGTGGTAAATGATGAAACACTTCTTTATGAGAAACA	240
Qy	262	GGAGTCCAGTTCCTCTTTTGGATGATACCTTCTCTTCTGCTCCCAATATGGAACATGTA	321
Db	241	GGAGTCCAGTTCCTCTTTTGGATGATACCTTCTCTTCTGCTCCCAATATGGAACATGTA	300
Qy	322	CTAGCAGTTGCCAATGAAGAAGCTTTGTTCGATTTGATTAACACAGATCACAAGTTTC	381
Db	301	CTAGCAGTTGCCAATGAAGAAGCTTTGTTCGATTTGATTAACACAGATCACAAGTTTC	360
Qy	382	AGAAAGAGTGTCTTCAAGAAATGGATGGCTCACTGGAAATGCCGTCTTTGACCTGGCTGG	441
Db	361	AGAAAGAGTGTCTTCAAGAAATGGATGGCTCACTGGAAATGCCGTCTTTGACCTGGCTGG	420
Qy	442	GTTCTCGGTGAATTTAACTTTGATACAGCAGCAGGTGATCAAAACAGCAAAATTTGGGAC	501
Db	421	GTTCTCGGTGAATTTAACTTTGATACAGCAGCAGGTGATCAAAACAGCAAAATTTGGGAC	480
Qy	502	GTAAAGCTGTGAGCTGATGAAATGCAAGAGTTCATCAATGACGCTCAAGTCAGTT	561
Db	481	GTAAAGCTGTGAGCTGATGAAATGCAAGAGTTCATCAATGACGCTCAAGTCAGTT	540
Qy	562	GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	621
Db	541	GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	600
Qy	622	GTCGTGGATACCAAGTGCACAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT	681
Db	601	GTCGTGGATACCAAGTGCACAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT	560
Qy	682	GGAGCTCACAAATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATTC	741
Db	661	GGAGCTCACAAATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATTC	720
Qy	742	AAAGGACTTGTCTCTCTGTGGATTTTCCAGCAAGTGTACTGTGCTCTTTTCAAGAC	801
Db	721	AAAGGACTTGTCTCTCTGTGGATTTTCCAGCAAGTGTACTGTGCTCTTTTCAAGAC	780
Qy	802	GAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAATATGGGATTTACGT	861
Db	781	GAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAATATGGGATTTACGT	840
Qy	862	AAGAATATACCTTATCGACAAGAACCCATAGCATCCAGTCTTCTGTACCCAGT	921
Db	841	AAGAATATACCTTATCGACAAGAACCCATAGCATCCAGTCTTCTGTACCCAGT	900
Qy	922	AGCAGCACTCGAAAACTTGGATATTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA	981

Db	901	AGCAGCACTCGAAAACTTGATATTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA	960
Qy	982	TTTGCTAAATGTCACAGACGATTAACATCTACATCTGTTTAAATATGACTGGGTGAAGACTTCT	1041
Db	961	TTTGCTAAATGTCACAGACGATTAACATCTAATGTTTAAATATGACTGGGTGAAGACTTCT	1020
Qy	1042	CCAGTGGCTATTCTCAATGGACACCAAGAACTCTACCTTTTATGTAATAATCCAGCCCTTAGT	1101
Db	1021	CCAGTGGCTATTCTCAATGGACACCAAGAACTCTACCTTTTATGTAATAATCCAGCCCTTAGT	1080
Qy	1102	CCAGATGACCCAGTTTGTAGTCAGTGGCTCAAGTGATGAAGCTGCCCTACATATGGAAGGTC	1161
Db	1081	CCAGATGACCCAGTTTGTAGTCAGTGGCTCAAGTGATGAAGCTGCCCTACATATGGAAGGTC	1140
Qy	1162	TCACACACCTGCGAACCTCTACTGCTGCTCTGGGTCATCTCAAGAGGTCAGCTCTGTG	1221
Db	1141	TCACACACCTGCGAACCTCTACTGCTGCTCTGGGTCATCTCAAGAGGTCAGCTCTGTG	1200
Qy	1222	TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACTCTGTCTGTGATGACAATACACTAAA	1281
Db	1201	TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACTCTGTCTGTGATGACAATACACTAAA	1260
Qy	1282	ATCTGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCAGGTC	1341
Db	1261	ATCTGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCAGGTC	1320
Qy	1342	GGTTGGGCTCTCAGAGAAGAAAAGAGTCAAGACCTGGCTAGTAAACAGTAAAGAGTAGC	1401
Db	1321	GGTTGGGCTCTCAGAGAAGAAAAGAGTCAAGACCTGGCTAGTAAACAGTAAAGAGTAGC	1380
Qy	1402	CAGAGTACTCTCGCAAGAGCCCGGTTAAAGTGAATCAATCAAACTTTTCCCGCTCA	1461
Db	1381	CAGAGTACTCTCGCAAGAGCCCGGTTAAAGTGAATCAATCAAACTTTTCCCGCTCA	1440
Qy	1462	TCCGAGCTTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATACTCTCAG	1521
Db	1441	TCCGAGCTTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATACTCTCAG	1500
Qy	1522	TTCTCTATTAAACCTCTCTGCGAAAGCCCGGCTCTCCCATCAACAGAGAGGCTCTGTCT	1581
Db	1501	TTCTCTATTAAACCTCTCTGCGAAAGCCCGGCTCTCCCATCAACAGAGAGGCTCTGTCT	1560
Qy	1582	TCCTCGCTCTCCCAAGCCACCTTCACTTTTCAAGATGTGCTATTAGAACTGGGTGACC	1641
Db	1561	TCCTCGCTCTCCCAAGCCACCTTCACTTTTCAAGATGTGCTATTAGAACTGGGTGACC	1620
Qy	1642	CGAACACCTTCTCATCACCCACCTCACTCCACCTGCTTCGGAGACCAAGATCATGTCT	1701
Db	1621	CGAACACCTTCTCATCACCCACCTCACTCCACCTGCTTCGGAGACCAAGATCATGTCT	1680
Qy	1702	CCGAGAAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCCAAGAGAGGCTTGTCTGAG	1761
Db	1681	CCGAGAAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCCAAGAGAGGCTTGTCTGAG	1740
Qy	1762	TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGAGAGTGTGAACAAAG	1821
Db	1741	TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGAGAGTGTGAACAAAG	1800
Qy	1822	TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTGATGGCCAAAGTTCTTCAATTTG	1881
Db	1801	TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTGATGGCCAAAGTTCTTCAATTTG	1860
Qy	1882	GATCTGTGCTGCTTGTGGTAAACAGGAAGACCTTTAGTAGGACTCTCTAGGCTCTACC	1941
Db	1861	GATCTGTGCTGCTTGTGGTAAACAGGAAGACCTTTAGTAGGACTCTCTAGGCTCTACC	1920
Qy	1942	AAATCAACAAATTAAGAGGAGCTGGTACAGATCTCTCAGAGCCCTCCCTCTCTCAGT	2001
Db	1921	AAATCAACAAATTAAGAGGAGCTGGTACAGATCTCTCAGAGCCCTCCCTCTCTCAGT	1980
Qy	2002	CCGTATGCTTCAGAAAGCTGTGGAACGCTACCTCTTCTTGTGAGACCTTGTGGAGAGGG	2061

1981 CCGTATGCTTCAGAAAGCTGTGAAGCGCTACCTCTCTCTTTGAGACCTTGTGGAGAGGG 2040
2062 TCTGAAATGGTATGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAAACTGGTGTGGCCATG 2121
2041 TCTGAAATGGTATGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAAACTGGTGTGGCCATG 2100
2122 GCAGCCAAACGGAGGCTGAGAAATCCATCTCCAGAAAGTCCGTCATCCAGACACCCCAAT 2181
2101 GCAGCCAAACGGAGGCTGAGAAATCCATCTCCAGAAAGTCCGTCATCCAGACACCCCAAT 2160
2182 TCAGGAGACAGAGCGGAAAGACATTCGCAAGCCCGGTCAACCATCAAGCCCGGTCCATG 2241
2161 TCAGGAGACAGAGCGGAAAGACATTCGCAAGCCCGGTCAACCATCAAGCCCGGTCCATG 2220
2242 AGGAAATCTGCACATATCTCCATAGAAAGTCCAGAGGAGCATTCCTGTGGTCCCTGAAACAC 2301
2221 AGGAAATCTGCACATATCTCCATAGAAAGTCCAGAGGAGCATTCCTGTGGTCCCTGAAACAC 2280
2302 TCACAGAAATATAGATTTCTAATCTGAGTGAGTACTGAGCTTTGGTCCACATAAACCAAG 2361
2281 TCACAGAAATATAGATTTCTAATCTGAGTGAGTACTGAGCTTTGGTCCACATAAACCAAG 2340
2362 CTGAGCTTTGGTCCATCAAAACAGATGAAATACAGAGTGACTCTATAACTCTGGTC 2421
2341 CTGAGCTTTGGTCCATCAAAACAGATGAAATACAGAGTGACTCTATAACTCTGGTC 2400
2422 TTAAAGAAAGCTGCCCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2481
2401 TTAAAGAAAGCTGCCCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2460
2482 TCTGTTCTCTACCATATATATATGACGCTCCGAGGATGATGCTGTGTAAATTT 2541
2461 TCTGTTCTCTACCATATATATATGACGCTCCGAGGATGATGCTGTGTAAATTT 2520
2542 TCATAAGTAAATTTCTGCTCTAGCATTTTGAATGAAATAGTCTTCACTTTTAAATAT 2601
2521 TCATAAGTAAATTTCTGCTCTAGCATTTTGAATGAAATAGTCTTCACTTTTAAATAT 2580
2602 TCATCTTCTCTATAATATGACATCCAGTTTCATGGAGCAAAAAACAAGTTTCTGTGTA 2661
2581 TCATCTTCTCTATAATATGACATCCAGTTTCATGGAGCAAAAAACAAGTTTCTGTGTA 2640
2662 TCCGTAAGCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGCCCTGTGAA 2721
2641 TCCGTAAGCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGCCCTGTGAA 2700
2722 GGCTGACTGAGAAATCCTCTGCTGGAAGACCCCTGGTTCTGTTCTGCTCCACATGTATA 2781
2701 GGCTGACTGAGAAATCCTCTGCTGGAAGACCCCTGGTTCTGTTCTGCTCCACATGTATA 2760
2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793

RESULT 12
US-10-848-755A-179
; Sequence 179, Application US/10848755A
; Publication No. US20050054826A1
; GENERAL INFORMATION:
; APPLICANT: Mao, Mao
; TITLE OF INVENTION: HUMAN DIAPHANOUS-3 GENE AND METHODS OF USE THEREFOR
; FILE REFERENCE: 9301-196-999
; CURRENT APPLICATION NUMBER: US/10/848,755A
; CURRENT FILING DATE: 2004-05-18
; PRIOR APPLICATION NUMBER: 60/471,842
; PRIOR FILING DATE: 2003-05-11
; NUMBER OF SEQ ID NOS: 275
; SOFTWARE: PatentIn version 3.2 CAM: 301891-999188
; SEQ ID NO 179
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-848-755A-179
Query Match 98.5%; Score 2789.8; DB 10; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 22 CGATAACGATTTGTGTGAGAGCGCAAGCTGCGATTTCTGCTGAACTTTGGAGGCAAT 81
DB 1 CGATAACGATTTGTGTGAGAGCGCAAGCTGCGATTTCTGCTGAACTTTGGAGGCAAT 60
QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGT 141
DB 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGT 120
QY 142 CTCGCGCAGCCCGAGCTTGGGCTCCTGAGAAATGAGTGGTCTTACAAATACCTCTTCAA 201
DB 121 CTCGCGCAGCCCGAGCTTGGGCTCCTGAGAAATGAGTGGTCTTACAAATACCTCTTCAA 180
QY 202 TCCCTTCTGACTGGTTATCAGTGAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA 261
DB 181 TCCCTTCTGACTGGTTATCAGTGAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA 240
QY 262 GGAGTCCCAGTTCCTCTCTTTTGGATGTAACCTTCTCTGCTCCCAATATGAAACATGTA 321
DB 241 GGAGTCCCAGTTCCTCTCTTTTGGATGTAACCTTCTCTGCTCCCAATATGAAACATGTA 300
QY 322 CTAGCAGTTGCCAATGAAGAGGCTTTGTTCGATTTGATATAACAGAAATCAAAAGTTTC 381
DB 301 CTAGCAGTTGCCAATGAAGAGGCTTTGTTCGATTTGATATAACAGAAATCAAAAGTTTC 360
QY 382 AGAAGAGTCTCTCAAGAAATGATGGCTCACTGGAATGCCCTTTTGACTCGGCCCTGG 441
DB 361 AGAAGAGTCTCTCAAGAAATGATGGCTCACTGGAATGCCCTTTTGACTCGGCCCTGG 420
QY 442 GTTCTGCTGAACTTAAACTTTGTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGAC 501
DB 421 GTTCTGCTGAACTTAAACTTTGTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGAC 480
QY 502 GTAAAGCTGGTGTGAGCTGATTTGGAACATGCAAGGTCAATCAATGAGCCTCAAGTCAGT 561
DB 481 GTAAAGCTGGTGTGAGCTGATTTGGAACATGCAAGGTCAATCAATGAGCCTCAAGTCAGT 540
QY 562 GCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGACGGGTGGAAGAGATGGCAACATTATG 621
DB 541 GCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGACGGGTGGAAGAGATGGCAACATTATG 600
QY 622 GTCTGGGATACAGGTGCACAAACAAAGATGGTTTTATAGCAAGTGAATCAAAATCAGT 681
DB 601 GTCTGGGATACAGGTGCACAAACAAAGATGGTTTTATAGCAAGTGAATCAAAATCAGT 660
QY 682 GGAGCTCAATACCTCTGAGACAGCAAAACCCCTTCAAAACCCAGAGAAACAGAAATTC 741
DB 661 GGAGCTCAATACCTCTGAGACAGCAAAACCCCTTCAAAACCCAGAGAAACAGAAATTC 720
QY 742 AAAGGACTTGTCTCTCTGAGATTTCCAGCAAAAGTGTACTGTGCTCTCTTTCAAGAC 801
DB 721 AAAGGACTTGTCTCTCTGAGATTTCCAGCAAAAGTGTACTGTGCTCTCTTTCAAGAC 780
QY 802 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAATATGGGATTTACGT 861
DB 781 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAATATGGGATTTACGT 840
QY 862 AAGAAATATCTGCTTATTCGACAAAGCCCATAGCATCAAGTCTTTCTGTACCCAGGT 921
DB 841 AAGAAATATCTGCTTATTCGACAAAGCCCATAGCATCAAGTCTTTCTGTACCCAGGT 900
QY 922 AGCAGCACTCGAAACACTTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
DB 901 AGCAGCACTCGAAACACTTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
QY 982 TTTGCTAATTCACAGACGATAACATCTACATGTTTAAATATGACCTGGGTGGAAGCTTCT 1041
DB 961 TTTGCTAATTCACAGACGATAACATCTACATGTTTAAATATGACCTGGGTGGAAGCTTCT 1020

1042 CCAGTGGCTATTTTCAATGGACACCGAACTCTACCTTTTATGTAAATCCAGCCTTAGT 1101
1021 CCAGTGGCTATTTTCAATGGACACCGAACTCTACCTTTTATGTAAATCCAGCCTTAGT 1080
1102 CCAGATGACAGATTTTATGTCAGTGGCTCAAGTGAAGAGCTGCCTACATATGGAAGGTC 1161
1081 CCAGATGACAGATTTTATGTCAGTGGCTCAAGTGAAGAGCTGCCTACATATGGAAGGTC 1140
1162 TCACACCTTGGCAACCTCTACTGTGCTCTGGGTCAATTTCAAGAGGTCAGTCTGTG 1221
1141 TCACACCTTGGCAACCTCTACTGTGCTCTGGGTCAATTTCAAGAGGTCAGTCTGTG 1200
1222 TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTTCTGATGACAATACACTAAAA 1281
1201 TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTTCTGATGACAATACACTAAA 1260
1282 ATCTGGCGCTTGAATAGAGCTTAGAGGAGAAACCGAGAGGTGATAACTTTTCCAGGTC 1341
1261 ATCTGGCGCTTGAATAGAGCTTAGAGGAGAAACCGAGAGGTGATAACTTTTCCAGGTC 1320
1342 GGTGGGCTCTCAGAGAAAGAGTCAAGACCTGGCTAGTAAACAGTAACAGTAGC 1401
1321 GGTGGGCTCTCAGAGAAAGAGTCAAGACCTGGCTAGTAAACAGTAACAGTAGC 1380
1402 CAGAGTACTCTGTCGAAGCCCGCAGGGTAAAGTGAATCCATCCCAATTCTTCCCGTCA 1461
1381 CAGAGTACTCTGTCGAAGCCCGCAGGGTAAAGTGAATCCATCCCAATTCTTCCCGTCA 1440
1462 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTTCCATCAATATCTCTACG 1521
1441 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTTCCATCAATATCTCTACG 1500
1522 TTCTCTATTAAGCTCTCTGCGAGGCCGCTCTCCATCAACAGAGAGGCTCTGTC 1581
1501 TTCTCTATTAAGCTCTCTGCGAGGCCGCTCTCCATCAACAGAGAGGCTCTGTC 1560
1582 TCCTCGTCTCTCCAGCCACCTTCAATCTTCAAGATGCGATTAGAACTGGTGACC 1641
1561 TCCTCGTCTCTCCAGCCACCTTCAATCTTCAAGATGCGATTAGAACTGGTGACC 1620
1642 CGAACACTTCTCATCACCAACCATCATCTCCACCTGCTTCCGAGACCAAGATCATGTCT 1701
1621 CGAACACTTCTCATCACCAACCATCATCTCCACCTGCTTCCGAGACCAAGATCATGTCT 1680
1702 CCAGAAAAGCCCTTATCTCTGTAGCCAGAGTATCCCAAGCAGAGGCTCTGTCGAG 1761
1681 CCAGAAAAGCCCTTATCTCTGTAGCCAGAGTATCCCAAGCAGAGGCTTCTCTGAG 1740
1762 TCTAGAAATAGATTAAGAGAGGCTAGACTCAAGCTGCTGGAGAGTGTGAACAAAAG 1821
1741 TCTAGAAATAGATTAAGAGAGGCTAGACTCAAGCTGCTGGAGAGTGTGAACAAAAG 1800
1822 TGTGTGAAGATTTGTAACCTGTGACTGAGCTTGTATGGCCCAAGTTGAAAATCTTCAATTTG 1881
1801 TGTGTGAAGATTTGTAACCTGTGACTGAGCTTGTATGGCCCAAGTTGAAAATCTTCAATTTG 1860
1882 GATCTGTGCTGCTTGTGTAAACAGGAAAGACTTATAGTAAGGACTCTCTAGTCTCTTACC 1941
1861 GATCTGTGCTGCTTGTGTAAACAGGAAAGACTTATAGTAAGGACTCTCTAGTCTCTTACC 1920
1942 AATCAAGCAAAATTGAAGAGCTGGTACCAGTATCTCAGAGCTCGTCTCTCTATCAGT 2001
1921 AATCAAGCAAAATTGAAGAGCTGGTACCAGTATCTCAGAGCTCGTCTCTCTATCAGT 1980
2002 CCGTATCTCTCAGAAAGCTGTGAACCGTACTACCTCTTCTTTGAGACCTTGTGGAAGGG 2061
1981 CCGTATCTCTCAGAAAGCTGTGAACCGTACTACCTCTTCTTTGAGACCTTGTGGAAGGG 2040
2062 TCTGAATGTGAGCAAGAGATAGTTCCCAAGAGATTAAGTGTGTTGGCCATG 2121
2041 TCTGAATGTGAGCAAGAGATAGTTCCCAAGAGATTAAGTGTGTTGGCCATG 2100

2122 GCAGCCAAACGGAAGGCTGAGAAATCCATCTCAAGAAGTCCGTATCCAGACACACCCAAT 2181
2101 GCAGCCAAACGGAAGGCTGAGAAATCCATCTCAAGAAGTCCGTATCCAGACACACCCAAT 2160
2182 TCAGAGAGACAGAGCGGAAAGACATTTGCAAGCCCGGTCAACATCAGCCGAGCTCCATG 2241
2161 TCAGAGAGACAGAGCGGAAAGACATTTGCAAGCCCGGTCAACATCAGCCGAGCTCCATG 2220
2242 AGGAAATCTGCACATACCTTCCATAGAAAGTCCAGAGAGACTTCTGTGCTCTGAAAC 2301
2221 AGGAAATCTGCACATACCTTCCATAGAAAGTCCAGAGAGACTTCTGTGCTCTGAAAC 2280
2302 TCAACAGAAATTAAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAAAA 2361
2281 TCAACAGAAATTAAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAAAA 2340
2362 CTGAGCTTTGGTCCACTAAAAAAGATGAAAAATACAGAGTGACTCTATTAATCTGGTC 2421
2341 CTGAGCTTTGGTCCACTAAAAAAGATGAAAAATACAGAGTGACTCTATTAATCTGGTC 2400
2422 TTTAGAAAGCTGCTTTTCTATTTTGTAGACAAATCTTTTCAACGCTGAAATGTACTAA 2481
2401 TTTAGAAAGCTGCTTTTCTATTTTGTAGACAAATCTTTTCAACGCTGAAATGTACTAA 2460
2482 TCTGGTTCTACTACCATTAATGTATATGACAGCTTCCGAGGATGAATGTGTTTAAAT 2541
2461 TCTGGTTCTACTACCATTAATGTATATGACAGCTTCCGAGGATGAATGTGTTTAAAT 2520
2542 TCATAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTA 2601
2521 TCATAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTA 2580
2602 TCATCTCTCTATATAATGACATCCAGTTCATGAGGCAAAACAAAGTTTCTTGTGA 2661
2581 TCATCTCTCTATATAATGACATCCAGTTCATGAGGCAAAACAAAGTTTCTTGTGA 2640
2662 TCTGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
2641 TCTGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
2722 GGCTGACTGAGAAATCTCTGCTGAGAGCCCTGTTCTGTTCTGCTCCACATGATATA 2781
2701 GGCTGACTGAGAAATCTCTGCTGAGAGCCCTGTTCTGTTCTGCTCCACATGATATA 2760
2782 ATTTTATTTGAATACATAATCTTTTCACTATG 2814
2761 ATTTTATTTGAATACATAATCTTTTCACTATG 2793

RESULT 13

US-10-330-773-673
; Sequence 673, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 673
; LENGTH: 4083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-673

Query Match 86.3%; Score 2442.4; DB 11; Length 4083;

Best Local Similarity 93.5%; Pred. No. 0;

Matches 2630; Conservative 0; Mismatches 11; Indels 171; Gaps 1;

3 CACGAGCGGAGTTGGAGCGGATAACGATTTGTGTTGTGAGAGCGGCAACGTCGATTC 62

37	Db	 CAGTGGCGGAGTTGGAGGGGATAACGATTTGTGTGTGAGAGCGCAAGCTGCGATTTC	96
63	Qy	 TGTCTGAATCTGGAGGCATTTCTACGACTTTCTCTCAGCTGAGGCTTTTCTCTCGGACCCT	122
97	Db	 TGTCTGAATCTGGAGGCATTTCTACGACTTTCTCTCAGCTGAGGCTTTTCTCTCGGACCCT	156
123	Qy	 GATGCTCTTCAATTCGGTGTCTCCGCCAGCCCCAGCTTTGGCGTCTCTGAGAAATGATGGTC	182
157	Db	 GATGCTCTTCAATTCGGTGTCTCCGCCAGCCCCAGCTTTGGCGTCTCTGAGAAATGATGGTC	216
183	Qy	 TTCACAATACCCCTCTTCAATCCCTTCTGACTGTGTTATCAGTGCAGTGGTAAATGATGAACA	242
217	Db	 TTCACAATACCCCTCTTCAATCCCTTCTGACTGTGTTATCAGTGCAGTGGTAAATGATGAACA	276
243	Qy	 CACTTCTTATGGAGAAACAGAGTCCCAAGTCTCTCTCTTTGGATGTACCTTCTCTCTGC	302
277	Db	 CACTTCTTATGGAGAAACAGAGTCCCAAGTCTCTCTCTCTTTGGATGTACCTTCTCTCTGC	336
303	Qy	 TCCCAATATGGAACATGTACTAGCAGTTGGCAATGGAAGAGGCTTCTGTCGATTTGATATAA	362
337	Db	 TCCCAATATGGAACATGTACTAGCAGTTGGCAATGGAAGAGGCTTCTGTCGATTTGATATAA	396
363	Qy	 CACAGAATCAAAAGTTTTCAGAAAGAAAGTGTCTTCAAAAGAAATGGATGGCTCACTGGGAATGC	422
397	Db	 CACAGAATCAAAAGTTTTCAGAAAGAAAGTGTCTTCAAAAGAAATGGATGGCTCACTGGGAATGC	456
423	Qy	 CGTCTTTGACCTGGCCCTGGGTTCTCTGGTGAACCTTAAACTTTGTTACAGCAGCAGGTGATCA	482
457	Db	 CGTCTTTGACCTGGCCCTGGGTTCTCTGGTGAACCTTAAACTTTGTTACAGCAGCAGGTGATCA	516
483	Qy	 AACAGCCAAATTTTGGGACGTAAAGAGCTGGTGAGCTGATTGGAAACATGCAAAAGTCAATCA	542
517	Db	 AACAGCCAAATTTTGGGACGTAAAGAGCTGGTGAGCTGATTGGAAACATGCAAAAGTCAATCA	576
543	Qy	 ATGCAGGCTCAAGTCAGTTGCTTTTCTAAAGTTTGGAGAAAGCTGTATTTCTGTACGGGTGG	602
577	Db	 ATGCAGGCTCAAGTCAGTTGCTTTTCTAAAGTTTGGAGAAAGCTGTATTTCTGTACGGGTGG	617
603	Qy	 AAGAGATGGCAACATTTATGTCTGGGATACAGGTGGCAACAAAGATGGGTTTTATAG	662
618	Db	 -----	617
663	Qy	 GCAAGTGAATCAAAATCAGTGGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCAAAACC	722
618	Db	 -----	617
723	Qy	 CAAGAAGAAAAGAAATCAAAAGGACTTGCTCTTCTGTGGATTTTCAGCAAAAGTGTATAC	782
618	Db	 -----GCTATCAGGAATTCAGCAAAAGTGTATAC	645
783	Qy	 TGTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATTAAT	842
646	Db	 TGTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATTAAT	705
843	Qy	 CAAAAGTATGGATTTAGCTAAGAAATATACCTTATTCGACAGAAACCCATAGCATCCAA	902
706	Db	 CAAAAGTATGGATTTAGCTAAGAAATATACCTTATTCGACAGAAACCCATAGCATCCAA	765
903	Qy	 GTCCTTCTGTACCCAGGTAGCAGCACTCGAAAACTTTGGATATTTCAAAGTCTGATTTTGGGA	962
766	Db	 GTCCTTCTGTACCCAGGTAGCAGCACTCGAAAACTTTGGATATTTCAAAGTCTGATTTTGGGA	825
963	Qy	 TTCACATGGCTCTACTTTATTTGCTAAATTTGACAGAGGATTAATATATATATATATATAT	1022
826	Db	 TTCACATGGCTCTACTTTATTTGCTAAATTTGACAGAGGATTAATATATATATATATATAT	885
1023	Qy	 GACTGGGTGGAAGACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTAA	1082
886	Db	 GACTGGGTGGAAGACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTAA	945
1083	Qy	 TGTAATAATCCAGCCCTTAGTCCAGATGACCAAGTCTTTTATAGTCAAGTGGCTCAAGTGAATG	1142

[illegible]

```
QY 2223 CATCAGCCGAGCTCCATGAGGAAATCTGCACATCTCCATAGAAAGTCCCGAGGGA 2282
DB 2086 CATCAGCCGAGCTCCATGAGGAAATCTGCACATCTCCATAGAAAGTCCCGAGGGA 2145
QY 2283 CTTCTGTGGTCTGAAACACTCAACAGAAATATAGATTTCTAATCTGAGTGATTTACTGAGC 2342
DB 2146 CTTCTGTGGTCTGAAACACTCAACAGAAATATAGATTTCTAATCTGAGTGATTTACTGAGC 2205
QY 2343 TTTGGTCCACTAAACCAAGCTGAGCTTTGGTCCACTAAACCAAGATGAAAAATACAGAG 2402
DB 2206 TTTGGTCCACTAAACCAAGCTGAGCTTTGGTCCACTAAACCAAGATGAAAAATACAGAG 2265
QY 2403 TGACTCTATAACTCTGTCTTTAAGAAAGCTGCTTTTCAATTTTAGACAAAAATCTTTTC 2462
DB 2266 TGACTCTATAACTCTGTCTTTAAGAAAGCTGCTTTTCAATTTTAGACAAAAATCTTTTC 2325
QY 2463 AACGCTGAATGTACCTTAATCTGTCTTACTTACCATTAATGTATATGACAGCTTCCCGAGGA 2522
DB 2326 AACGCTGAATGTACCTTAATCTGTCTTACTTACCATTAATGTATATGACAGCTTCCCGAGGA 2385
QY 2523 TGAATGCTGTGTTAAATTTTATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAG 2582
DB 2386 TGAATGCTGTGTTAAATTTTATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAG 2445
QY 2583 TCTTCACTTTTAAATTTATTTATCTCTCTATAATATGATCCAGATTTTATGGAGGCA 2642
DB 2446 TCTTCACTTTTAAATTTATTTATCTCTCTATAATATGATCCAGATTTTATGGAGGCA 2505
QY 2643 AAAAAACAAGTTCTTGTATCTGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702
DB 2506 AAAAAACAAGTTCTTGTATCTGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2565
QY 2703 ACAGCATGAGGCTGTGAAAGGCTGACTGAGAAATCTCTGCTGAAAGCCCTGGTCTGT 2762
DB 2566 ACAGCATGAGGCTGTGAAAGGCTGACTGAGAAATCTCTGCTGAAAGCCCTGGTCTGT 2625
QY 2763 TCTGCTCTCAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTATG 2814
DB 2626 TCTGCTCTCAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTATG 2677
```

RESULT 14

```
US-10-221-625-131
; Sequence 131, Application US/10221625
; Publication No. US20040033942A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: LU, Dying Aina M.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Valda
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: MATHUR, Preete
; APPLICANT: SHAH, Purvi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: REDDY, Roopa
; TITLE OF INVENTION: TRANSCRIPTION FACTORS
; FILE REFERENCE: PF-0761 PCT
; CURRENT APPLICATION NUMBER: US/10/221,625
; NUMBER FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PERL Program
; SEQ ID NO 131
; LENGTH: 1871
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
OTHER INFORMATION: Incyte ID No. US20040033942A1 1377380CB1
US-10-221-625-131
```

```
Query Match 53.7%; Score 1520; DB 8; Length 1871;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1284 CTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGAGGTGATAAACTTTCCACGCTGGG 1343
DB 1 CTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGAGGTGATAAACTTTCCACGCTGGG 60
QY 1344 TTGGGCTCTCAGAGAAAAAGAGTCAAGACTGCGCTTAGTAAACAGTAAACAGTAAAC 1403
DB 61 TTGGGCTCTCAGAGAAAAAGAGTCAAGACTGCGCTTAGTAAACAGTAAACAGTAAAC 120
QY 1404 GAGTACTCTGCGCAAGCCCGGAGGTAAAGTCAATCCATCCATTTCTCCCGTCATC 1463
DB 121 GAGTACTCTGCGCAAGCCCGGAGGTAAAGTCAATCCATCCATTTCTCCCGTCATC 180
QY 1464 CGCAGCTTGTGCCCCAAGCTGTGCGAGACCTCCCTCTTCCCTTCAAAATACCTACGTT 1523
DB 181 CGCAGCTTGTGCCCCAAGCTGTGCGAGACCTCCCTCTTCCCTTCAAAATACCTACGTT 240
QY 1524 CTCTATTAAAACTCTCTGCGCAAGCCCGGCTTCCCATCAAAGAGGCTCTGCTCTC 1583
DB 241 CTCTATTAAAACTCTCTGCGCAAGCCCGGCTTCCCATCAAAGAGGCTCTGCTCTC 300
QY 1584 CTCGCTCTCTCCAGCCACCTTCACTTCAAGATGTCGATTAGAACTGGGTGA -CCC 1642
DB 301 CTCGCTCTCTCCAGCCACCTTCACTTCAAGATGTCGATTAGAACTGGGTGA -CCCC 360
QY 1643 GAACACTTCTCATCACCACCATCCTCCTGCGAGACCAAGATCATGCTCTC 1702
DB 361 GAACACTTCTCATCACCACCATCCTCCTGCGAGACCAAGATCATGCTCTC 420
QY 1703 CGAGAAAAAGCCCTTATCTGTCGAGCAGAACTCATCCCAAGCAGAGGCTTGTCTGAGT 1762
DB 421 CGAGAAAAAGCCCTTATCTGTCGAGCAGAACTCATCCCAAGCAGAGGCTTGTCTGAGT 480
QY 1763 CTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGTCGAGAGTGGAACAAAGT 1822
DB 481 CTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGTCGAGAGTGGAACAAAGT 540
QY 1823 GTGTGAAGAGTTGTAACCTGTGTGACTGTGAGCTTGTGAGCCAAAGTGTGAAATCTTCA 1882
DB 541 GTGTGAAGAGTTGTAACCTGTGTGACTGTGAGCTTGTGAGCCAAAGTGTGAAATCTTCA 600
QY 1883 ATCTGTGCTGCTTGTGCTGTAACCCAGGAAGACCTTAGTAAAGGACTCTTAGGTCTCTAC 1942
DB 601 ATCTGTGCTGCTTGTGCTGTAACCCAGGAAGACCTTAGTAAAGGACTCTTAGGTCTCTAC 660
QY 1943 AATCAAGCAAAATTAAGAGGCTGGTACAGTATCTCAGAGGCTCGCTCTCTATCAGTC 2002
DB 661 AATCAAGCAAAATTAAGAGGAGCTGGTACAGTATCTCAGAGGCTCGCTCTCTATCAGTC 720
QY 2003 CGTATGCTTCAAGAGCTGTGGAACGCTACCTCTCTTTTGGAGACCTTGTGAGAGGCT 2062
DB 721 CGTATGCTTCAAGAGCTGTGGAACGCTACCTCTCTTTTGGAGACCTTGTGAGAGGCT 780
QY 2063 CTGAAATGGTAGGCAAGAGAAATAGTTTCCCGAGAGAAATAAAAGTGTGTTGGCCATGG 2122
DB 781 CTGAAATGGTAGGCAAGAGAAATAGTTTCCCGAGAGAAATAAAAGTGTGTTGGCCATGG 840
QY 2123 CAGCCAAACGGAAGGCTGAGAAATCCATCTCAGAGAGTCCGTCATCCCGAGACACCAAT 2182
DB 841 CAGCCAAACGGAAGGCTGAGAAATCCATCTCAGAGAGTCCGTCATCCCGAGACACCAAT 900
QY 2183 CCAGGAGACAGAGCGGAAAGACATTGCCAAGCCCGCTCACCATCAGCCAGCTCCATCA 2242
DB 901 CCAGGAGACAGAGCGGAAAGACATTGCCAAGCCCGCTCACCATCAGCCAGCTCCATCA 960
QY 2243 GGAAATCTGCACATCTTCCATAGAAAGTCCCGAGGAGCTTCTGTGCTCTCTGACACT 2302
```

961	GGAAAACTCGGACATATCTCCATAGAAAGTCCCGAGGAGCACTTCTGTGGCTCGGAACACT	1020
2303	CAACAGAAATTATAGATTTCTAATCTGAGTGAAGTTACTGAGCTTTGGTCCACTAAAACAAGC	2362
1021	CAACAGAAATTATAGATTTCTAATCTGAGTGAAGTTACTGAGCTTTGGTCCACTAAAACAAGC	1080
2363	TGAGCTTTGGTCCACTAAAACAAGATGAAAAATACAAGAGTGACTCTATAACTCTGGTCT	2422
1081	TGAGCTTTGGTCCACTAAAACAAGATGAAAAATACAAGAGTGACTCTATAACTCTGGTCT	1140
2423	TTAAGAAAGCTGCCCTTTTCATTTTTTAGCAAAATCTTTTCAACGCTGAAATGTACCTAAT	2482
1141	TTAAGAAAGCTGCCCTTTTCATTTTTTAGCAAAATCTTTTCAACGCTGAAATGTACCTAAT	1200
2483	CTGTTCTTACTACCATAATGTATATGACAGCTTCCGAGGATGAATGCTGTGTTTAAATTT	2542
1201	CTGTTCTTACTACCATAATGTATATGACAGCTTCCGAGGATGAATGCTGTGTTTAAATTT	1260
2543	CATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTAAT	2602
1261	CATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTAAT	1320
2603	CATCTTCTCTATAATAATGACATCCCAAGTTTCATGGAGGCAAAAAACAAGTTTCTTGTAT	2662
1321	CATCTTCTCTATAATAATGACATCCCAAGTTTCATGGAGGCAAAAAACAAGTTTCTTGTAT	1380
2663	CCTGAAACTTTCTATGTCTCAGTGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAAG	2722
1381	CCTGAAACTTTCTATGTCTCAGTGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAAG	1440
2723	GCTGACTGAGAAATCTCTGCTGAAGACCCCTGGTTCTGTGTTCTGCCCTCCCAACATGTATAA	2782
1441	GCTGACTGAGAAATCTCTGCTGAAGACCCCTGGTTCTGTGTTCTGCCCTCCCAACATGTATAA	1500
2783	TTTTATTGAAATACATAAATCTTTTCACTATG	2814
1501	TTTTATTGAAATACATAAATCTTTTCACTATG	1532

```

RESULT 15
US-10-330-773-670
; Sequence 670, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FIELD OF INVENTION: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 670
; LENGTH: 2426
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-330-773-670

```

Qy	768	CCAGCAAGAGTGTACTGTGGTCTCTTTTCAAGACGAGAACTA	827
Db	809	CCAGCAGAGTGTACTGTGGTCTCTTTTCAAGATGAGAA	868
Qy	828	TGTGGATGGGATAATCAAAAGTATGGGATTTACGTAAGA	887
Db	869	CGTGGATGGNATNATCAAAAGTATGGGATTTGGCGAAGA	928
Qy	888	ACCCATAGCAATCAAGTCTTTTCTGTATCCAGGTAGCAG	947
Db	929	ACCCATAGCAATCAAGTCTTTTCTGTATCCAGGTAGCAG	988
Qy	948	AAGTCTGATTTTGGATTTCCACTGGCTCTACTTTATTTG	1007
Db	989	GAGTTTTGGTTTTAGACTCTACTGGCTCTACTTTATTTG	1048
Qy	1008	CTACATGTTTAAATATGACTGGGTTGAAGACTTCTCCAG	1067
Db	1049	CTATATGTTCAATATGACTGGCTTTAAGACTTCTCCGGT	1108
Qy	1068	GAACTCTACCTTTATGTATAAATCCAGCTTGTAGTCAG	1127
Db	1109	GAACTCTACCTTTTATGTATAAATCAAGTCTTAGTCAG	1168
Qy	1128	CTCAAGTGATGAAGCTGCCTACATATATGGAAGTCTCC	1187
Db	1169	TTCAAGTGATGAAGCTGCCTACATATTTGGAAGTTCCT	1228
Qy	1188	GCTCCTGGGTCAATCTCAAGAGTCAAGTCTGTGTGTCT	1247
Db	1229	GCTCCTGGGTCAATCTCAAGAGTCAAGTCTGTGTGTCT	1288
Qy	1248	GATTGTACCTGTTCTGTATGACAAATACACTAAAAATC	1307
Db	1289	GATTGCAACCTGCTCTGTATGATATACACTGAAANATC	1348
Qy	1308	GGAGAAACAGGAGGTGATAAACTTTCCACGGTGGGTTG	1367
Db	1349	GGAGAAACCTTTCCAGTATGATTTCCATAGTGGTGGAC	1405
Qy	1368	GTCGAAGCTGGCTCTAGTAAACAGTAAAGGATAGCAG	1427
Db	1406	AGTGAAAGCTGCGCCAGTAAACGGTAAACAGTACGAG	1465
Qy	1428	GGTAAAGTGCAATCCATCCAAATTTCTCCCGTCAATCG	1487
Db	1466	AGCCAGACGAGTCCATCCATCTCTCTTGTGTAGCAGCT	1525
Qy	1488	TGGAGACCTCCCTCTTCTTCAAAATACTCTTAGTTCCT	1547
Db	1526	AGGAGACCTCCCTCTTCTTCAAGTACCCCCCAATTTCT	1585
Qy	1548	GGCCCCGTCTCCATCAACGAGAGNGCTCTGTCTCTCT	1607
Db	1586	GACCCGTCTTCTAGTTCAGCAGAGAGGCTCCATCTCT	1645
Qy	1608	ATCTTTCAAGATGTGGATTAGAACTGGGTGAGCCGGAAC	1667
Db	1646	ATCTTTCAAGATGTGGTTAGAAACTGGGTGACCCGGAAC	1705
Qy	1668	CATCTCACTGCTTTCGGAGACCAAGATCATGTCTCCGA	1727
Db	1706	CACCTCCACCTGCTTCTCGAGCAAAAGATCTCATCTCC	1765
Qy	1728	CCGAAAGTCATCCCAACGACGAGCTCTCTGAGTCTFAG	1787
Db	1766	CCGAAAGTCATCAACGACGAGTCTCTCTCTGATCTFAG	1825
Qy	1788	AGACTCAAGCTGTCTGGAGAGTGTCAAAACAAAAAGTGT	1847
Db	1826	TGACTCAAGCTGTCTGGAGAGTGTCAAAACAAAAAGTGT	1885

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 05:58:58 ; Search time 518 Seconds
(without alignments)
10226.079 Million cell updates/sec

Title: US-10-726-160-1

Perfect score: 2831

Sequence: 1 ggcacgagcggttgag.....atgaaaaaaaaaaaaaaaaa 2831

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /EMC Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:*
- 2: /EMC Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:*
- 3: /EMC Celerra_SIDS3/ptodata/2/ina/6A COMB.seq:*
- 4: /EMC Celerra_SIDS3/ptodata/2/ina/6B COMB.seq:*
- 5: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:*
- 6: /EMC Celerra_SIDS3/ptodata/2/ina/H COMB.seq:*
- 7: /EMC Celerra_SIDS3/ptodata/2/ina/pCTUS COMB.seq:*
- 8: /EMC Celerra_SIDS3/ptodata/2/ina/pp COMB.seq:*
- 9: /EMC Celerra_SIDS3/ptodata/2/ina/RE COMB.seq:*
- 10: /EMC Celerra_SIDS3/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500.4	17.7	815	4	US-09-297-648-3050 Sequence 3050, Ap
2	421.4	14.9	477	4	US-09-297-648-5115 Sequence 5115, Ap
3	300	10.6	300	4	US-09-297-648-99 Sequence 99, Appl
4	300	10.6	300	4	US-09-297-648-1050 Sequence 1050, Ap
5	290.4	10.3	346	4	US-09-880-107-3634 Sequence 3634, Ap
6	276.8	9.8	290	3	US-09-513-999C-21911 Sequence 21911, A
7	156	5.5	413	3	US-09-513-999C-21244 Sequence 21244, A
8	53	1.9	7218	2	US-08-232-463-14 Sequence 14, Appl
9	48.2	1.7	1548	3	US-09-614-221A-136 Sequence 136, App
10	44.2	1.6	7218	2	US-08-232-463-14 Sequence 14, Appl
11	44	1.6	832	3	US-09-621-976-2813 Sequence 2813, Ap
12	42.8	1.5	1141	3	US-09-806-708B-22 Sequence 22, Appl
13	40.8	1.4	2025	3	US-09-252-991A-1536 Sequence 1536, Ap
14	40.6	1.4	77851	3	US-09-949-016-12508 Sequence 12508, A
15	40.6	1.4	77857	3	US-09-949-016-13211 Sequence 13211, A
16	40.6	1.4	77857	3	US-09-949-016-13212 Sequence 13212, A
17	40.6	1.4	77940	3	US-09-949-016-12509 Sequence 12509, A
18	40.4	1.4	84525	3	US-09-949-016-16578 Sequence 16578, A
19	40.2	1.4	1141	3	US-09-806-708B-22 Sequence 22, Appl
20	40	1.4	601	3	US-09-949-016-174701 Sequence 174701, A
21	39.6	1.4	63658	3	US-09-949-016-13238 Sequence 13238, A
22	39.6	1.4	64489	3	US-09-949-016-11766 Sequence 11766, A
23	38.8	1.4	700	3	US-09-735-271-383 Sequence 383, App

Sequence 6297, Ap
Sequence 2, Appli
Sequence 2, Appli
Sequence 996, App
Sequence 3156, Ap
Sequence 18438, A
Sequence 1433, Ap
Sequence 16480, A
Sequence 211, App
Sequence 36, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 1960, Ap
Sequence 266, App
Sequence 4891, Ap
Sequence 897, App

24 38.8 1.4 1593 3 US-09-248-796A-6297
25 38.6 1.4 1211 3 US-09-063-743-2
26 38.6 1.4 1211 3 US-09-590-540-2
27 38 1.3 2724 4 US-10-094-749-996
28 37.8 1.3 561 3 US-09-270-767-3156
29 37.8 1.3 561 3 US-09-270-767-18438
30 37.8 1.3 1230 3 US-09-252-991A-1433
31 37.6 1.3 225127 3 US-09-949-016-16480
32 37.6 1.3 1082144 4 US-09-531-120-211
33 37.4 1.3 475 3 US-09-702-705-36
34 37.4 1.3 475 3 US-09-736-457-36
35 37.4 1.3 475 3 US-09-614-124B-36
36 37.4 1.3 475 3 US-09-671-325-36
37 37.4 1.3 475 3 US-09-589-184-36
38 37.4 1.3 475 3 US-09-658-824-36
39 37.4 1.3 475 3 US-10-017-754-36
40 37.4 1.3 475 3 US-09-651-563-36
41 37.4 1.3 475 3 US-09-519-643-36
42 37.4 1.3 744 3 US-09-248-796A-1960
43 37.4 1.3 1262 3 US-09-149-476-266
44 37.4 1.3 1431 3 US-09-248-796A-4891
45 37.4 1.3 6158 3 US-09-799-451-897

ALIGNMENTS

RESULT 1

US-09-297-648-3050
; Sequence 3050, Application US/09297648
; Patent No. 6964868
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime A.
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominiguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassan, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, William Lee
; APPLICANT: Stache-Crain, Birjit
; TITLE OF INVENTION: No. 6964868 Human Genes and Gene Expression
; FILE REFERENCE: 2300-1481
; CURRENT APPLICATION NUMBER: US/09/297,648
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,666
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,515
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 5252
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3050
; LENGTH: 815
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(815)
; OTHER INFORMATION: n = A,T,C or G
US-09-297-648-3050

Query Match 17.7%; Score 500.4; DB 4; Length 815;
Best Local Similarity 96.9%; Pred. No. 1.1e-139;
Matches 559; Conservative 0; Mismatches 13; Indels 5; Gaps 5;

QY 1780 AGGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAGTGTGTGAAGAGTTGTAAAC 1839
Db 65 ACGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAGTGTGTGAAGAGTTGTAAAC 124
QY 1840 TGTGTGACTGAGCTTGATGCCCAAGTTGAAATCTTCATTTGGATCTGTGCTTGGCT 1899
Db 125 TGTGTGACTGAGCTTGATGCCCAAGTTGAAATCTTCATTTGGATCTGTGCTTGGCT 184
QY 1900 GGTAAACAGAGACCTTAGTAAGGACTCTCTAGGTCTTACCAATCAAGCAAAATTTGAA 1959
Db 185 GGTAAACAGAGACCTTAGTAAGGACTCTCTAGGTCTTACCAATCAAGCAAAATTTGAA 244
QY 1960 GGAGCTGGTACCAGTATCTCAGAGCTCCGTCTCTCTATCAGTCCGTATGCTTCAGAAAGC 2019
Db 245 GGAGCTGGTACCAGTATCTCAGAGCTCCGTCTCTCTATCAGTCCGTATGCTTCAGAAAGC 304
QY 2020 TGTGGACCTACCTCTCTCTTTGAGACCTTGTGAGAGAGGTCTGAAATGTAGGCATA 2079
Db 305 TGTGGACCTACCTCTCTCTTTGAGACCTTGTGAGAGAGGTCTGAAATGTAGGCATA 364
QY 2080 GAGATAGTTTCCACAGAGATAAAACTGTTTGTGGCCATGCGACCAACCGGAAGGCT 2139
Db 365 GAGATAGTTTCCACAGAGATAAAACTGTTTGTGGCCATGCGACCAACCGGAAGGCT 424
QY 2140 GAGATCCATCTCCAGAGTCCGTATCCAGACACCCAAATTCAGGAGACAGAGCGGA 2199
Db 425 GAGATCCATCTCCAGAGTCCGTATCCAGACACCCAAATTCAGGAGACAGAGCGGA 484
QY 2200 AAGACATTCGCAAGCCCGGTACCATCAGCCCGAGCTCATGAGGAAATCTGCACATAC 2259
Db 485 AAGACATTCGCAAGCCCGG - CACCATCAGCCCGAGCTTCATGAGGAAATCTGCACATAC 543
QY 2260 TTCCATAGAAAGTCCACAGGAGAC - TTCTGTGTCTCTGAACACTCAA - CAGAAATTATA - G 2316
Db 544 TTCCATAGAAAGTCCACAGGAGAC - TTCTGTGTCTCTGAACACTCAA - CAGAAATTATA - G 603
QY 2317 ATTCTAATCTGAG - TGAGTTACTGAGCTTTTGGTCCAC 2352
Db 604 ATTCTAATCTGAGTTGAGTTACTGAGCTTTTGGTCCC 640

RESULT 2

US-09-297-648-5115
; Sequence 5115, Application US/09297648
; Patent No. 6964868
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominiguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassan, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje

; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, William Lee
; APPLICANT: Stache-Crain, Birjit
; TITLE OF INVENTION: No. 6964868el Human Genes and Gene Expression
; FILE REFERENCE: 2300-1481
; CURRENT APPLICATION NUMBER: US/09/297,648
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,666
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,515
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 5252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5115
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-297-648-5115

Query Match 14.9%; Score 421.4; DB 4; Length 477;
Best Local Similarity 99.5%; Pred. No. 4.7e-116;
Matches 433; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1785 GCTAGACTCAAGCTGTCTGGAGAGTGTGTAACAAAGTGTGTGAAGAGTTGTAACTGTGT 1844
Db 1 GCTAGACTCAAGCTGTCTGGAGAGTGTGTAACAAAGTGTGTGAAGAGTTGTAACTGTGT 60
QY 1845 GACTGAGCTTCATGCGCCCAAGTTGAAATCTTCATTTGGATCTGTCTGCTGCTGCTGTA 1904
Db 61 GACTGAGCTTCATGCGCCCAAGTTGAAATCTTCATTTGGATCTGTCTGCTGCTGCTGTA 120
QY 1905 CCAGGAAGACCTTAGTAAGGACTCTCTAGGTCTTACCAAAATCAAGCAAAATTTGAAGGAGC 1964
Db 121 CCAGGAAGACCTTAGTAAGGACTCTCTAGGTCTTACCAAAATCAAGCAAAATTTGAAGGAGC 180
QY 1965 TGGTACCAGTATCTCAGAGCTCCGTCTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGG 2024
Db 181 TGGTACCAGTATCTCAGAGCTCCGTCTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGG 240
QY 2025 AACGCTACCTCTTCCCTTTGAGACCTTGTGGAGAGAGGTCTGAAATGTTAGGCAAGAGAA 2084
Db 241 AACGCTACCTCTTCCCTTTGAGACCTTGTGGAGAGAGGTCTGAAATGTTAGGCAAGAGAA 300
QY 2085 TAGTTTCCCGAGAGATAAAAACTGGTGTGTGGCCATGGCAGCAACCGAAGGCTGAGAA 2144
Db 301 TAGTTTCCCGAGAGATAAAAACTGGTGTGTGGCCATGGCAGCAACCGAAGGCTGAGAA 359
QY 2145 TCCATCTCAGAAAGTCCGTCTATCCAGACACCCCAATTTCCAGGAGACAGAGCGGAAAGNC 2204
Db 360 TCCATCTCAGAAAGTCCGTCTATCCAGACACCCCAATTTCCAGGAGACAGAGCGGAAAGNC 419
QY 2205 ATTGCCAAGCCCGGT 2219
Db 420 ATTGCCAAGCCCGCT 434

RESULT 3

US-09-297-648-99

```
; Sequence 99, Application US/09297648
; Patent No. 6964868
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassan, Altaf
; APPLICANT: Lamsen, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Dickson, Mark
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, William Lee
; APPLICANT: Stache-Crain, Birjit
; TITLE OF INVENTION: No. 6964868el Human Genes and Gene Expression
; FILE REFERENCE: 2300-1481
; CURRENT APPLICATION NUMBER: US/09/297,648
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,666
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,515
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/105,234
; NUMBER OF SEQ ID NOS: 5252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-297-648-99

Query Match      10.6%; Score 300; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.2e-79;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1785 GCTAGACTCAAGCTGCTGAGAGTGTGAACAAAAGTGTGAAGAGTTGTAACCTGTGT 1844
Db 1 GCTAGACTCAAGCTGCTGAGAGTGTGAACAAAAGTGTGAAGAGTTGTAACCTGTGT 60

QY 1845 GACTGAGCTTGATGGCCAAAGTTGAAAATCTTCAATTTGGATCTGTGCTGCTTGTCTGTTAA 1904
Db 61 GACTGAGCTTGATGGCCAAAGTTGAAAATCTTCAATTTGGATCTGTGCTGCTTGTCTGTTAA 120

QY 1905 CCAGGAAGACCTTAGTAAGACTCTCTTAGTCTCTACCAATCAAGCAAAATTTGAAGGAGC 1964
Db 121 CCAGGAAGACCTTAGTAAGACTCTCTTAGTCTCTACCAATCAAGCAAAATTTGAAGGAGC 180

QY 1965 TGGTACCAAGTATCTAGAGGCTCGTCTCTATCAGTCCGTATGCTTTCAGAAAGCTGTGS 2024
Db 181 TGGTACCAAGTATCTAGAGGCTCGTCTCTATCAGTCCGTATGCTTTCAGAAAGCTGTGS 240

QY 2025 AACGCTACCTCTTCCCTTTGAGACCTTGTGGAGAGGGTCTGAAATGTTAGGCAAGAGAA 2084
Db 241 AACGCTACCTCTTCCCTTTGAGACCTTGTGGAGAGGGTCTGAAATGTTAGGCAAGAGAA 300

RESULT 4
US-09-297-648-1050
; Sequence 1050, Application US/09297648
; Patent No. 6964868
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassan, Altaf
; APPLICANT: Lamsen, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Dickson, Mark
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, William Lee
; APPLICANT: Stache-Crain, Birjit
; TITLE OF INVENTION: No. 6964868el Human Genes and Gene Expression
; FILE REFERENCE: 2300-1481
; CURRENT APPLICATION NUMBER: US/09/297,648
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,666
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,515
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/105,234
; NUMBER OF SEQ ID NOS: 5252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1050
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-297-648-1050

Query Match      10.6%; Score 300; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.2e-79;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1632 CTGGGTGACCCGAAACACCTTCTCTATCACCACCCATCACTCCACTGTCTCGAGACCAA 1691
Db 1 CTGGGTGACCCGAAACACCTTCTCTATCACCACCCATCACTCCACTGTCTCGAGACCAA 60

QY 1692 GATCATGTCTCGAGAAAAGCCCTTATTCCTGTGAGCCAGAGTCAATCCCAAGCAGAGGC 1751
Db 61 GATCATGTCTCGAGAAAAGCCCTTATTCCTGTGAGCCAGAGTCAATCCCAAGCAGAGGC 120

QY 1752 TTGCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTCGAGAGTGT 1811
Db 121 TTGCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTCGAGAGTGT 180

QY 1812 GAAACAAAAGTGTGTAAGAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGA 1871
Db 181 GAAACAAAAGTGTGTAAGAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGA 240
```


